



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 162063

TO: Sumesh Kaushal  
Location: 2B85/2C70  
Art Unit: 1633  
Friday, August 19, 2005

Case Serial Number: 10/617835

From: Deirdre Arnold  
Location: Biotech-Chem Library  
REM 1A64  
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

### Search Notes

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold



**This Page Blank (uspro)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2005, 21:01:19 ; Search time 126 Seconds  
(without alignments)  
982.249 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645

Sequence: 1 MRARLLIPLRSVFILSAAG.....SHEGYGSDAVRHRGQGP 320

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16dec04:\*

1: \_geneseq1980s:\*\n2: geneseqp1990s:\*\n3: geneseqp2000s:\*\n4: geneseqp2001s:\*\n5: geneseqp2002s:\*\n6: geneseqp2003as:\*\n7: geneseqp2003bs:\*\n8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1645	100.0	320	2	AAW18785
2	1641	99.8	320	2	AAJ38540
3	1641	99.8	320	3	AAJ75812
4	1641	99.8	320	3	AAJ74926
5	1641	99.8	320	3	AAJ25657
6	1641	99.8	320	3	AAJ58587
7	1641	99.8	320	6	ABP76947
8	1634	99.3	320	2	AAJ38538
9	1634	99.3	320	3	AAJ74927
10	1634	99.3	320	3	AAJ75811
11	1634	99.3	320	3	AAJ25656
12	1634	99.3	320	3	AAJ58586
13	1634	98.7	320	2	AAJ38539
14	1624	98.7	320	3	AAJ74928
15	1624	98.7	320	3	AAJ75813
16	1624	98.7	320	3	AAJ25658
17	1624	98.7	320	3	AAJ58588
18	1300	79.0	263	6	ABP79829
19	1042	63.3	213	2	AAJ38537
20	1030	62.6	313	2	AAJ38537
21	1030	62.6	313	6	ABP78726
22	1030	62.6	313	6	ABP80805
23	1025.5	62.3	314	2	AAJ38538
24	1022	62.1	313	2	AAJ38538
25	650	39.5	197	2	AAJ38538

26	117.5	7.1	1250	4	ABJ65508
27	117.5	7.1	1250	4	ABJ65470
28	106.5	6.5	190	7	ADP04155
29	106.5	6.5	506	7	ABJ66253
30	102.5	6.2	1250	2	AAJ59361
31	101	6.1	445	6	ABJ24747
32	100.5	6.1	1499	5	ABJ58409
33	99.5	6.0	756	5	ABJ58421
34	99	6.0	872	6	ABJ27543
35	97	5.9	515	7	ABJ64118
36	96.5	5.9	685	3	AAJ76449
37	96	5.8	274	8	ADJ58765
38	96	5.8	274	8	ADJ58769
39	96	5.8	684	3	AAJ59318
40	96	5.8	729	4	ABJ29904
41	95	5.8	274	8	ADJ58761
42	95	5.8	274	8	ADJ58763
43	95	5.8	274	8	ADJ58766
44	95	5.8	424	6	ADJ07548
45	94.5	5.7	343	2	AAJ00088

## ALIGNMENTS

RESULT 1	AAW18785	standard; protein; 320 AA.
ID	AAW18785	
XX	AAW18785;	
AC	XX	
XX	XX	
DT	10-DEC-1997	(first entry)
XX	XX	
DE	Neisseria adhesion protein (from Orf1).	
XX	XX	
KW	Neisseria gonorrhoeae; adhesion; lipoprotein; OrfA; OrfI; OrfB.	
XX	XX	
OS	Neisseria gonorrhoeae.	
XX	XX	
PN	DE19534579-A1.	
XX	XX	
PD	20-MAR-1997.	
XX	XX	
PF	18-SEP-1995; 95DE-01034579.	
XX	XX	
PR	18-SEP-1995; 95DE-01034579.	
XX	XX	
PA	(PLAC ) MAX PLANCK GBS FOERDERUNG WISSENSCHAFTEN.	
XX	XX	
DR	WPI; 1997-180942/17.	
XX	XX	
PT	Nucleic acids encoding Neisseria adhesion proteins - for therapeutic and diagnostic use.	
XX	XX	
PS	Claim 29; Page 11-13; 20pp; German.	
XX	XX	
CC	OrfA and OrfB in complexes with the protein PilC are capable of adhering to human cells. Products obtained from the DNA are useful in medicaments, CC diagnostic compans. and vaccines, esp. for treatment of Neisseria CC gonorrhoea and N. meningitidis infections	
XX	XX	
SO	Sequence 320 AA;	
Query Match	100.0%; Score 1645; DB 2; Length 320;	
Best Local Similarity	100.0%; Pred. No. 1.2e-151;	
Matches	320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MRARLLIPLRSVFILSAAGTGTGTPSHGGGKFAVEQELVVAASRAAVKMDLQALGR 60	
DB	1 MRARLLIPLRSVFILSAAGTGTGTPSHGGGKFAVEQELVVAASRAAVKMDLQALGR 60	
QY	61 KVALYATMGDSGSLTGRYSIDALIRGEYINSPAVRTDYTPYETTAETTSGLTG 120	

```

Db      61 KVALIATMDGSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGLTG 120
QY      121 LTTSLSTLNAPALSRTOGSGSRSSLGINIGMGDYRNETLTTPRPTAFLSHVQVVF 180
Db      121 LTTSLSTLNAPALSRTOGSGSRSSLGINIGMGDYRNETLTTPRPTAFLSHVQVVF 180
QY      181 FLRGIDVVSPPANADTDVFINIDVFCTIRNRTEMLYNAETLKAQTKLEYFAVDRTNNKLL 240
Db      181 FLRGIDVVSPPANADTDVFINIDVFCTIRNRTEMLYNAETLKAQTKLEYFAVDRTNNKLL 240
QY      241 IKPRTNAFEAAKYKENYALMGPYKVSKGIKPTGELMVDPSDIRPYGNHTGNSAPSV EADN 300
Db      241 IKPRTNAFEAAKYKENYALMGPYKVSKGIKPTGELMVDPSDIRPYGNHTGNSAPSV EADN 300
QY      301 SHEGYGYSDEAVRQHRQGP 320
Db      301 SHEGYGYSDEAVRQHRQGP 320

RESULT 2
AA38540
ID      AA38540 standard; protein; 320 AA.
XX      AAY38540;
XX      08-OCT-1999 (first entry)
XX      DE      Neisseria gonorrhoeae antigenic protein encoded by ORF15.
XX      KM      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX      KW      treatment; Neisseria infection; meningitis; septicemia; gonorrhea.
XX      OS      Neisseria gonorrhoeae.
XX      PN      WO9924578-A2.
XX      PD      20-MAY-1999.
XX      PF      09-OCT-1998; 98WO-IB001665.
XX      PR      06-NOV-1997; 97GB-00023516.
XX      PR      14-NOV-1997; 97GB-00024130.
XX      PR      18-NOV-1997; 97GB-00024386.
XX      PR      27-NOV-1997; 97GB-00025158.
XX      PR      10-DEC-1997; 97GB-00026147.
XX      PR      14-JAN-1998; 98GB-00000759.
XX      PR      01-SEP-1998; 98GB-00019016.
XX      PA      (CHIR-) CHIRON SPA.
XX      PI      Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX      WPI: 1999-327407/27.
XX      DR      N-PSDB; AA212008.
XX      PT      Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX      diagnosis, treatment and prevention of infection.
XX      PS      Claim 4; Page 100; 524pp; English.
XX      CC      Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis and
XX      CC      N. gonorrhoeae antigenic proteins. They are encoded by open reading
XX      CC      frames (ORFs) AA211972-212358. The antigenic proteins, their fragments,
XX      CC      their nucleic acids and antibodies are used for diagnosis, prevention (as
XX      CC      vaccines) or treatment of Neisseria infections, such as meningitis,
XX      CC      septicemia and gonorrhea. Both organisms are closely related. Fragments
XX      CC      of the nucleic acids are useful as hybridisation probes and antisense
XX      CC      reagents
XX      SQ      Sequence 320 AA;
Query Match      99.8%; Score 1641; DB 2; Length 320;
Best Local Similarity      99.7%; Pred. No. 2.8e-151;

```

```

Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY      1 MRARLLPILESVLTILACGTLTGIPSHGGKRFABEOLVAASABAAYKMDLQALHGR 60
Db      1 MRARLLPILESVLTILACGTLTGIPSHGGKRFABEOLVAASABAAYKMDLQALHGR 60
QY      61 KVALIATMDGSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGLTG 120
Db      61 KVALIATMDGSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGLTG 120
QY      121 LTTSLSTLNAPALSRTOGSGSRSSLGINIGMGDYRNETLTTPRPTAFLSHVQVVF 180
Db      121 LTTSLSTLNAPALSRTOGSGSRSSLGINIGMGDYRNETLTTPRPTAFLSHVQVVF 180
QY      181 FLRGIDVVSPPANADTDVFINIDVFCTIRNRTEMLYNAETLKAQTKLEYFAVDRTNNKLL 240
Db      181 FLRGIDVVSPPANADTDVFINIDVFCTIRNRTEMLYNAETLKAQTKLEYFAVDRTNNKLL 240
QY      241 IKPRTNAFEAAKYKENYALMGPYKVSKGIKPTGELMVDPSDIRPYGNHTGNSAPSV EADN 300
Db      241 IKPRTNAFEAAKYKENYALMGPYKVSKGIKPTGELMVDPSDIRPYGNHTGNSAPSV EADN 300
QY      301 SHEGYGYSDEAVRQHRQGP 320
Db      301 SHEGYGYSDEAVRQHRQGP 320

RESULT 3
AA75812
ID      AA75812 standard; protein; 320 AA.
XX      AAY75812;
XX      DT      12-SEP-2003 (revised)
XX      DT      21-MAR-2000 (first entry)
XX      DE      Neisseria gonorrhoeae ORF 406 protein sequence SEQ ID NO:3112.
XX      KM      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX      KW      antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX      KW      antibacterial; gene therapy.
XX      OS      Neisseria gonorrhoeae.
XX      PN      WO9957280-A2.
XX      PD      11-NOV-1999.
XX      PF      30-APR-1999; 99WO-US009346.
XX      PR      01-MAY-1998; 98US-0083758P.
XX      PR      31-JUL-1998; 98US-0094869P.
XX      PR      02-SEP-1998; 98US-0098994P.
XX      PR      02-SEP-1998; 98US-0099062P.
XX      PR      09-OCT-1998; 98US-0103749P.
XX      PR      09-OCT-1998; 98US-0103794P.
XX      PR      09-OCT-1998; 98US-0103796P.
XX      PR      25-FEB-1999; 99US-0121528P.
XX      PA      (CHIR ) CHIRON CORP.
XX      PI      (GENO-) INST GENOMIC RES.
XX      PI      Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
XX      PI      Petersen J, Pizza M, Rappuoli R, Ratti G, Scarlato E, Scarselli M;
XX      PI      Tettelin H, Venter JC;
XX      DR      WPI: 2000-062150/05.
XX      DR      N-PSDB; AA254614.
XX      PT      Novel Neisserial polypeptides predicted to be useful antigens for
XX      PT      vaccines and diagnostics.
XX      PS      Example 1, Page 109; 1453pp; English.

```



```

XX AA253015 to AA254536, AA254577 to AA254615, and AA774253 to AA775941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254617 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to Neisseria
CC bacteria (e.g. meningitis and septicemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
CC field)
XX
SQ Sequence 320 AA;
Query Match 99.8%; Score 1641; DB 3; Length 320;
Best Local Similarity 99.7%; Pred. No. 2.8e-151;
Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRARLLIPILPSVFLISACGTLTGIPSHGGKRFVBEQELVAASARAANKMDLQALHGR 60
DB 1 MRARLLIPILPSVFLISACGTLTGIPSHGGKRFVBEQELVAASARAANKMDLQALHGR 60
QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGGTLG 120
DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGGTLG 120
QY 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
DB 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
QY 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
DB 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
QY 181 FLRGIDVVS PANADTVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
DB 181 FLRGIDVVS PANADTVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
QY 241 IKPKTNAEFAAYKENYALMWPYKSKIKPTGELMVPFSDIQPYGNHTGNSAPVEADN 300
DB 241 IKPKTNAEFAAYKENYALMWPYKSKIKPTGELMVPFSDIQPYGNHTGNSAPVEADN 300
QY 301 SHEGYGSDAEVROHROGQP 320
DB 301 SHEGYGSDAEVROHROGQP 320

RESULT 4
AA774926 standard; protein, 320 AA.
XX AA774926;
XX
XX 12-SEP-2003 (revised)
XX 21-MAR-2000 (first entry)
DE Neisseria gonorrhoeae ORF 406 protein sequence SEQ ID NO:1326.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
XX Neisseria gonorrhoeae.
OS
XX WO9957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
XX 31-JUL-1998; 98US-0094869P.
XX 02-SEP-1998; 98US-0098994P.

```

```

PR 02-SEP-1998; 98US-0099062P.
PR 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
PR 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
XX (GENO-) CHIRON CORP.
XX PA (INST GENOMIC RES.
XX
XX Frazer C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarcelll M;
XX Tetzelin H, Venter JC;
XX
XX WPI: 2000-062150/05.
XX N-PSDB; AA253688.
XX
XX Novel Neisseria polypeptides predicted to be useful antigens for
XX vaccines and diagnostics.
XX
XX Claim 2; Page 728; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AA774253 to AA775941
XX represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX and polypeptides. AA254537 to AA254576 and AA254616 to AA254617 represent
XX PCR primers used in the exemplification of the present invention. The
XX polypeptides, the polynucleotides, antibodies and compositions of the
XX invention can be used as vaccines, as diagnostic reagents, and as
XX immunogenic compositions. The polypeptides can be used in the manufacture
XX of medicaments for treating or preventing infection due to Neisseria
XX bacteria (e.g. meningitis and septicemia), to detect the presence of
XX Neisseria bacteria, or to raise antibodies. They may also be used to
XX screen for agonists or antagonists, which may themselves have use as
XX antibacterial agents. The polynucleotides of the invention may also be
XX used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
XX field)
XX
SQ Sequence 320 AA;
Query Match 99.8%; Score 1641; DB 3; Length 320;
Best Local Similarity 99.7%; Pred. No. 2.8e-151;
Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRARLLIPILPSVFLISACGTLTGIPSHGGKRFVBEQELVAASARAANKMDLQALHGR 60
DB 1 MRARLLIPILPSVFLISACGTLTGIPSHGGKRFVBEQELVAASARAANKMDLQALHGR 60
QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGGTLG 120
DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYYPRYETTAETTSGGTLG 120
QY 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
DB 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
QY 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
DB 121 LTTSLSTINAPALSTGDSGSRSLGNTGMDYNNETLTNPRTAFLSHVQTVF 180
QY 181 FLRGIDVVS PANADTVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
DB 181 FLRGIDVVS PANADTVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
QY 241 IKPKTNAEFAAYKENYALMWPYKSKIKPTGELMVPFSDIQPYGNHTGNSAPVEADN 300
DB 241 IKPKTNAEFAAYKENYALMWPYKSKIKPTGELMVPFSDIQPYGNHTGNSAPVEADN 300
QY 301 SHEGYGSDAEVROHROGQP 320
DB 301 SHEGYGSDAEVROHROGQP 320

RESULT 5
AAB25657 standard; protein, 320 AA.
XX AAB25657;
XX

```

DT 12-SEP-2003 (revised)  
DT 04-DEC-2000 (first entry)  
XX  
DE N. gonorrhoeae amino acid sequence g406.pep SEQ ID NO:1037.  
XX  
KM Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KM Meningococcus B; MenB.  
XX  
OS Neisseria gonorrhoeae.  
XX  
EN WO200022430-A2.  
XX  
PD 20-APR-2000.  
XX  
PF 08-OCT-1999; 99WO-US023573.  
XX  
PR 09-OCT-1998; 98US-0103794P.  
PR 30-APR-1999; 99US-0132068P.  
XX  
XX (CHIR ) CHIRON CORP.  
PI Frazer CM, Hickey E, Peterson J, Tettein H, Venter JC;  
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappuoli R, Pizza M;  
XX  
DR WPI: 2000-318079/27.  
DR N-PSDB; AAF81297.  
XX  
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used  
PT in the diagnosis and treatment of N. meningitidis infection and other  
PT Neisserial infections, for example, N.gonorrhoea.  
XX  
PS Claim 14; Page 108; 1760pp; English.  
XX  
XX The present invention describes methods of obtaining immunogenic proteins  
CC from Neisseria genomic sequences. AAF81453 to AAF82414 represent  
CC specifically claimed Neisseria meningitidis genomic DNA sequences;  
CC AAF81260 to AAF81303 and AAF825620 to AAF825663 represent Neisseria DNA  
CC sequences and their corresponding proteins; AAF81254 to AAF81259 and  
CC AAF81304 to AAF81321 represent PCR primers used in the isolation of  
CC Neisseria meningitidis DNA sequences; and AAF81322 to AAF81452 represent  
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all  
CC used in the exemplification of the present invention. The nucleic acid  
CC sequences, protein sequences, and antibodies against them, can be used in  
CC the manufacture of a composition. The composition can be used as a  
CC medicament (or in the manufacture of a medicament) for treating  
CC preventing or diagnosing infection due to Neisserial bacteria. For  
CC example, some of the identified proteins could be components of vaccines  
CC against Meningococcus B; against all serotypes; and/or against all  
CC pathogenic Neisseriae. Identification of sequences from the bacterium  
CC will also facilitate production of biological probes, particularly  
CC organism-specific probes. Attempts to make efficacious Meningococcus B  
CC vaccines have failed mainly due to antigen tolerance. Multivalent  
CC vaccines have also been tried but none have successfully overcome  
CC antigenic variability. The provision of further, complete sequences may  
CC provide an opportunity to identify secreted or surface exposed proteins  
CC that may be presumed targets for the immune system and which are not  
CC antigenically variable or at least more conserved than other more  
CC variable regions. (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 320 AA;  
Query Match 99.8%; Score 1641; DB 3; Length 320;  
Best Local Similarity 99.7%; Pred. No. 2.8e-151; Indels 0; Gaps 0;  
Matches 319; Conservative 1; Mismatches 0;  
QY 1 MRARLLIPILSFVILSACGTLTGIPSHGGKRFPAVEOELVAASRAAVKMDIQALHGR 60  
DB 1 MRARLLIPILSFVILSACGTLTGIPSHGGKRFPAVEOELVAASRAAVKMDIQALHGR 60  
QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSFAVRDYYPRYETTAETTSGLITG 120

DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSFAVRDYYPRYETTAETTSGLITG 120  
QY 121 LTTSLSTLNAPALSRTOGSDGSRSSLGILNIGMGDYNRETLTTPRDTAFSLHVOYVF 180  
DB 121 LTTSLSTLNAPALSRTOGSDGSRSSLGILNIGMGDYNRETLTTPRDTAFSLHVOYVF 180  
QY 181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTENHLVNAETLKAOTKLEYFVADRTNKKL 240  
DB 181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTENHLVNAETLKAOTKLEYFVADRTNKKL 240  
QY 241 IKPRTNAFEAAVKENYVLMGPKYVSKGIPRTESGLMDPSPDIRPGNHTGNSAPVEADN 300  
DB 241 IKPRTNAFEAAVKENYVLMGPKYVSKGIPRTESGLMDPSPDIRPGNHTGNSAPVEADN 300  
QY 301 SHEGYSDEAVRQHRGQGP 320  
DB 301 SHEGYSDEAVRQHRGQGP 320  
RESULT 6  
AAB58587  
ID AAB58587 standard; protein; 320 AA.  
AC AAB58587;  
XX  
DT 12-SEP-2003 (revised)  
DT 13-MAR-2001 (first entry)  
XX  
XX N. gonorrhoeae amino acid sequence g406.pep SEQ ID NO:77.  
XX  
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KM diagnosis; antigen; detection; infection; gene therapy; antibacterial.  
XX  
OS Neisseria gonorrhoeae.  
XX  
XX WO200006791-A1.  
EN  
XX 09-NOV-2000.  
PD  
XX 08-MAR-2000; 2000WO-US005928.  
PF  
XX 30-APR-1999; 99US-0132068P.  
PR 08-OCT-1999; 99WO-US023573.  
PR 28-FEB-2000; 2000GB-00004695.  
XX  
XX (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
PI Pizza M, Hickey E, Peterson J, Tettein H, Venter JC,  
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappuoli R, Frazer CM, Grandi G;  
XX  
DR WPI: 2000-647603/62.  
DR N-PSDB; AAF21582.  
XX  
PT Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections.  
XX  
PS Example 1; Page 109; 692pp; English.  
XX  
XX The present invention describes the full length genome of Neisseria  
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF22613  
CC represent fragments of the NMB genomic sequence, as the sequence was too  
CC long to go in a record on its own it was split into 8 sequences which  
CC overlap each other at the beginning and end of each sequence by 49980 bp  
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of  
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of  
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins  
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR  
CC primers which are used in the exemplification of the present invention.  
CC The NMB genome and fragments from it have antibacterial activity, and can  
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
CC and/or antibodies which binds to the proteins can be used in compositions

CC for treating or preventing infection due to *Neisseria* bacteria or as a  
 CC diagnostic reagent for detecting the presence of *Neisseria* bacteria or  
 CC of antibodies raised to *Neisseria* bacteria. Computers, computer memory,  
 CC computer storage medium or computer databases can be used in a search to  
 CC identify open reading frames (ORFs) or coding sequences within the NMB  
 CC genome. The DNA sequences provide further opportunities to find antigenic  
 CC or immunogenic proteins which are more effective in vaccines (Updated on  
 CC 12-SEP-2003 to standardise OS field)

XX  
 SQ Sequence 320 AA;

Query Match 99.8%; Score 1641; DB 3; Length 320;  
 Best Local Similarity 99.7%; Pred. No. 2.8e-151;  
 Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRARLLIPILFSVFLISACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGR 60  
 DB 1 MRARLLIPILFSVFLISACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGR 60  
 QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGTLG 120  
 DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGTLG 120  
 QY 121 LTTSLSTINAPALSRSTQSDGSRSLGLNIGMGDYRNETLTPRDTAFSLHVQYVF 180  
 DB 121 LTTSLSTINAPALSRSTQSDGSRSLGLNIGMGDYRNETLTPRDTAFSLHVQYVF 180  
 QY 181 FLRGIDVVS PANADTDVFINIDVFETIRNRTEMHLYNAETLKAQTKLEFAVDRTNKLL 240  
 DB 181 FLRGIDVVS PANADTDVFINIDVFETIRNRTEMHLYNAETLKAQTKLEFAVDRTNKLL 240  
 QY 241 IKPKTNAEFAAYKENYALMGPYKYSKGIKPTGELMVPFSDIRPYGNHTGNSAPSEADN 300  
 DB 241 IKPKTNAEFAAYKENYALMGPYKYSKGIKPTGELMVPFSDIRPYGNHTGNSAPSEADN 300  
 QY 301 SHEGYGSDAEAVRHRQGQP 320  
 DB 301 SHEGYGSDAEAVRHRQGQP 320

RESULT 7  
 ABP76947

ID ABP76947 standard; protein; 320 AA.

XX AC ABP76947;

XX DT 07-MAR-2003 (first entry)

XX DE N. gonorrhoeae amino acid sequence SEQ ID 424.

XX KW Antibacterial; infection; vaccine; gene therapy.

XX OS *Neisseria gonorrhoeae*.

XX PN WO200279243-A2.

XX PD 10-OCT-2002.

XX PF 12-FEB-2002; 2002WO-IB002069.

XX PR 12-FEB-2001; 2001GB-00003424.

XX PA (CHIR-) CHIRON SPA.

XX PI Fontana M, Pizsa M, Massignani V, Monaci E;

XX DR MPI: 2003-058415/05.

XX DR N-PSDB; AB237917.

PT New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a  
 PT medicament for treating or preventing N. gonorrhoeae infection.

XX PS Disclosure; Page 220; 815pp; English.

XX The present invention relates to proteins from *Neisseria gonorrhoeae*.  
 CC Also disclosed are the nucleic acid molecules encoding the proteins and  
 CC antibodies that specifically bind to the proteins. The composition  
 CC comprising the protein, nucleic acid or antibody is useful for the  
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
 CC infection, this may be in the form of a vaccine or gene therapy.  
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid  
 CC molecules of the invention

XX  
 SQ Sequence 320 AA;

Query Match 99.8%; Score 1641; DB 6; Length 320;  
 Best Local Similarity 99.7%; Pred. No. 2.8e-151;  
 Matches 319; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRARLLIPILFSVFLISACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGR 60  
 DB 1 MRARLLIPILFSVFLISACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGR 60  
 QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGTLG 120  
 DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGTLG 120  
 QY 121 LTTSLSTINAPALSRSTQSDGSRSLGLNIGMGDYRNETLTPRDTAFSLHVQYVF 180  
 DB 121 LTTSLSTINAPALSRSTQSDGSRSLGLNIGMGDYRNETLTPRDTAFSLHVQYVF 180  
 QY 181 FLRGIDVVS PANADTDVFINIDVFETIRNRTEMHLYNAETLKAQTKLEFAVDRTNKLL 240  
 DB 181 FLRGIDVVS PANADTDVFINIDVFETIRNRTEMHLYNAETLKAQTKLEFAVDRTNKLL 240  
 QY 241 IKPKTNAEFAAYKENYALMGPYKYSKGIKPTGELMVPFSDIRPYGNHTGNSAPSEADN 300  
 DB 241 IKPKTNAEFAAYKENYALMGPYKYSKGIKPTGELMVPFSDIRPYGNHTGNSAPSEADN 300  
 QY 301 SHEGYGSDAEAVRHRQGQP 320  
 DB 301 SHEGYGSDAEAVRHRQGQP 320

RESULT 8  
 AAY38538

ID AAY38538 standard; protein; 320 AA.

XX AC AAY38538;

XX DT 08-OCT-1999 (first entry)

XX DE *Neisseria meningitidis* strain B antigen encoded by ORF15.

XX KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KW treatment; *Neisseria* infection; meningitis; septicemia; gonorrhea.

XX OS *Neisseria meningitidis*.

XX PN WO9924578-A2.

XX PD 20-MAY-1999.

XX PF 09-OCT-1998; 98WO-IB001665.

XX PR 06-NOV-1997; 97GB-00023516.

XX PR 14-NOV-1997; 97GB-00024190.

XX PR 18-NOV-1997; 97GB-00024386.

XX PR 27-NOV-1997; 97GB-00025158.

XX PR 10-DEC-1997; 97GB-00026147.

XX PR 14-JAN-1998; 98GB-00000759.

XX PR 01-SEP-1998; 98GB-00019016.

XX PA (CHIR-) CHIRON SPA.

XX Massignani V, Rappuoli R, Pizsa M, Scarlato V, Grandi G;

XX WP1, 1999-327407/27.  
 DR N-PSDB; AAZ12006.  
 XX  
 PT Proteins from *Neisseria meningitidis* and *N. gonorrhoeae* useful for  
 diagnosis, treatment and prevention of infection.  
 XX  
 PS Claim 4; Page 98; 524pp; English.  
 XX  
 CC Amino acid sequences AAY38499-Y38944 represent *Neisseria meningitidis* and  
 CC *N. gonorrhoeae* antigenic proteins. They are encoded by open reading  
 CC frames (ORFs) AAZ11972-212358. The antigenic proteins, their fragments,  
 CC their nucleic acids and antibodies are used for diagnosis, prevention (as  
 CC vaccines) or treatment of *Neisseria* infections, such as meningitis,  
 CC septicemia and gonorrhea. Both organisms are closely related. Fragments  
 CC of the nucleic acids are useful as hybridisation probes and antisense  
 CC reagents  
 XX  
 SQ Sequence 320 AA;  
 Query Match 99.3%; Score 1634; DB 2; Length 320;  
 Best Local Similarity 99.1%; Pred. No. 1.4e-150;  
 Matches 317; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MRARLLIPILFSVFILSACGTLTGIPSHGGGKRFVQEGLVAASARAAYKMDLQALHGR 60  
 DB 1 MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFVQEGLVAASARAAYKMDLQALHGR 60  
 QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTG 120  
 DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTG 120  
 QY 121 LTTSLSTLNAPALSRTOSSDGSRSLSGLNIGMGDYRNELTTNPRDTAFSLHVOTVF 180  
 DB 121 LTTSLSTLNAPALSRTOSSDGSRSLSGLNIGMGDYRNELTTNPRDTAFSLHVOTVF 180  
 QY 181 FLRGIDVSPANADTDVFINIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
 DB 181 FLRGIDVSPANADTDVFINIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
 QY 241 IKPRTNAFEAAYKENYALMWGPYKVSKGIRPTEGLMVDPSDIRPYGNHTGNSAPSEADN 300  
 DB 241 IKPRTNAFEAAYKENYALMWGPYKVSKGIRPTEGLMVDPSDIRPYGNHTGNSAPSEADN 300  
 QY 301 SHEGYSDEAVRQHRQGP 320  
 DB 301 SHEGYSDEAVRQHRQGP 320  
 RESULT 9  
 AAY74927  
 ID AAY74927 standard; protein; 320 AA.  
 XX  
 AC AAY74927;  
 XX  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE *Neisseria meningitidis* ORF 406 protein sequence SEQ ID NO:1328.  
 XX  
 KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy.  
 XX  
 OS *Neisseria meningitidis*.  
 XX  
 PN MO9957280-A2.  
 PD 11-NOV-1999.  
 PF 30-APR-1999; 99WO-US009346.  
 PR 01-MAY-1998; 98US-0083758P.  
 PR 31-JUL-1998; 98US-0094869P.

PR 02-SEP-1998; 98US-0098994P.  
 PR 02-SEP-1998; 98US-0099062P.  
 PR 09-OCT-1998; 98US-0103749P.  
 PR 09-OCT-1998; 98US-0103794P.  
 PR 09-OCT-1998; 98US-0103796P.  
 PR 25-FEB-1999; 99US-0121528P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M,  
 PI Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M,  
 PI Tettein H, Venter JC;  
 XX  
 DR WP1: 2000-062150/05.  
 DR N-PSDB; AAZ53689.  
 XX  
 PT Novel *Neisseria* polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics.  
 XX  
 PS Claim 2; Page 728-729; 1453pp; English.  
 XX  
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of the  
 CC invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the manufacture  
 CC of medicaments for treating or preventing infection due to *Neisseria*  
 CC bacteria (e.g. meningitis and septicemia), to detect the presence of  
 CC *Neisseria* bacteria, or to raise antibodies. They may also be used to  
 CC screen for agonists or antagonists, which may themselves have use as  
 CC antibacterial agents. The polynucleotides of the invention may also be  
 CC used in gene therapy protocols  
 XX  
 SQ Sequence 320 AA;  
 Query Match 99.3%; Score 1634; DB 3; Length 320;  
 Best Local Similarity 99.1%; Pred. No. 1.4e-150;  
 Matches 317; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MRARLLIPILFSVFILSACGTLTGIPSHGGGKRFVQEGLVAASARAAYKMDLQALHGR 60  
 DB 1 MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFVQEGLVAASARAAYKMDLQALHGR 60  
 QY 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTG 120  
 DB 61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTG 120  
 QY 121 LTTSLSTLNAPALSRTOSSDGSRSLSGLNIGMGDYRNELTTNPRDTAFSLHVOTVF 180  
 DB 121 LTTSLSTLNAPALSRTOSSDGSRSLSGLNIGMGDYRNELTTNPRDTAFSLHVOTVF 180  
 QY 181 FLRGIDVSPANADTDVFINIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
 DB 181 FLRGIDVSPANADTDVFINIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
 QY 241 IKPRTNAFEAAYKENYALMWGPYKVSKGIRPTEGLMVDPSDIRPYGNHTGNSAPSEADN 300  
 DB 241 IKPRTNAFEAAYKENYALMWGPYKVSKGIRPTEGLMVDPSDIRPYGNHTGNSAPSEADN 300  
 QY 301 SHEGYSDEAVRQHRQGP 320  
 DB 301 SHEGYSDEAVRQHRQGP 320  
 RESULT 10  
 AAY75811  
 ID AAY75811 standard; protein; 320 AA.  
 XX  
 AC AAY75811;  
 XX

DT	21-MAR-2000	(first entry)
XX	Neisseria meningitidis ORF 406 protein sequence SEQ ID NO:3110.	
XX	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;	
KW	antibacterial; gene therapy.	
XX	Neisseria meningitidis.	
OS	WO957280-A2.	
XX	11-NOV-1999.	
PD	30-APR-1999; 99WO-US009346.	
PF	01-MAY-1998; 98US-0083758P.	
XX	31-JUL-1998; 98US-0094869P.	
PR	02-SEP-1998; 98US-0098994P.	
PR	02-SEP-1998; 98US-0099062P.	
PR	09-OCT-1998; 98US-0103749P.	
PR	09-OCT-1998; 98US-0103794P.	
PR	09-OCT-1998; 98US-0103796P.	
PR	25-FEB-1999; 99US-0121528P.	
XX	(CHIR) CHIRON CORP.	
PA	(GENO) INST GENOMIC RES.	
XX	Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;	
PI	Petersen J, Pizzi M, Rappuoli R, Ratti G, Scalato E, Scarfelli M;	
PI	Tetelli H, Venter JC;	
XX	WPI; 2000-062150/05.	
DR	N-PSDB; AA254613.	
XX	Novel Neisserial polypeptides predicted to be useful antigens for	
PT	vaccines and diagnostics.	
XX	Example 1; Page 109; 1453pp; English.	
XX	AA253015 to AA254536, AA254577 to AA254615, and AA774253 to AA775941	
CC	represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides	
CC	and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent	
CC	PCR primers used in the exemplification of the present invention. The	
CC	polypeptides, the polynucleotides, antibodies and compositions of the	
CC	invention can be used as vaccines, as diagnostic reagents, and as	
CC	immunogenic compositions. The polypeptides can be used in the manufacture	
CC	of medicaments for treating or preventing infection due to Neisserial	
CC	bacteria (e.g. meningitis and septicemia), to detect the presence of	
CC	Neisseria bacteria, or to raise antibodies. They may also be used to	
CC	screen for agonists or antagonists, which may themselves have use as	
CC	antibacterial agents. The polynucleotides of the invention may also be	
CC	used in gene therapy protocols	
XX	Sequence 320 AA;	
SO		
QY	Query Match 99.3%; Score 1634; DB 3; Length 320;	
DB	Best local similarity 99.1%; Pred. No. 1.4e-150;	
DB	Matches 317; Conservative 2; Mismatches 1; Indels 0; Gaps 0	
QY	1 MRARLLDILFVSFLLSACGTLTGIPSHGGKRFVBOELVAASRAAAYKMDLQALHGR 60	
DB	1 MQARLLDILFVSFLLSACGTLTGIPSHGGKRFVBOELVAASRAAAYKMDLQALHGR 60	
QY	61 KVALYIATMGDGGSSLTGGRYSIDALIRGEYINSPAVRTDVTTPRYETTAETSGGLTG 120	
DB	61 KVALYIATMGDGGSSLTGGRYSIDALIRGEYINSPAVRTDVTTPRYETTAETSGGLTG 120	
QY	121 LTTSLSTINAPALSTRQSDGSSRSLSGLNIGMGDYRNETLTTPRODPAFLSHLYQVTF 180	
DB	121 LTTSLSTINAPALSTRQSDGSSRSLSGLNIGMGDYRNETLTTPRODPAFLSHLYQVTF 180	
QY	181 FLRGIDVVSANADTVFINIDVFGITRNRTEMLHYNATLKAQTKLEFVADRTRNKKLL 240	

Dd		 FLRGIDVPSPAADTDFINVIDFGIIRKRTMHLNAAETLAKAQYLBFAVDRIKKLL 240
Oy	241 IKPKTNAEBAAYKENVALMMGPGKYKSQGIKPTEGMLVPDSDIRPYNHTGNAPSVEADN 300	
Dd	241 IKPKTNABEAPAAKENVYALMMGPYKYSKGIKPTFGLMWDPDSDIRPYNHNGNSAPSVEADN 300	
Oy	301 SHEGYGSDEAVRQRQCQP 320	
Dd	301 SHEGYGSDEAVRQRQCQP 320	
RESULT 11		
AAB25656		
ID	AAB25656 standard; protein; 320 AA.	
XX AC AAB25656;		
XX XX		
DT 04-DEC-2000	(first entry)	
XX DE N. meningitidis amino acid sequence m406.pgp SEQ ID NO:1035.		
XX KM Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;		
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;		
KM Meningococcus B; MenB.		
XX OS Neisseria meningitidis .		
XX PN WO200022430-A2.		
PD 20-APR-2000.		
PF 08-OCT-1999;	99WO-USO23573.	
XX PR 09-OCT-1998;	98US--0103794P.	
PR 30-APR-1999;	99US--0132068P.	
XX PA (CHIR ) CHIRON CORP.		
P1 Frizer GM, Hickey E, Peterson J, Tetrelin H, Venter JC;		
P1 Masingham V, Galeotti C, Mora M, Ratti G, Scarelli M, Scarlato V;		
P1 Pappucci R, Pizza M,		
DR WPI; 2000-318079/27.		
N-PSTDB; AAA81296.		
Isolated nucleotide sequences of Neisseria meningitidis which can be used in the diagnosis and treatment of N. meningitidis infection and other Neisserial infections, for example, N.gonorrhoea.		
Claim 14; Page 108; 1760pp; English.		
The present invention describes methods of obtaining immunogenic proteins from Neisseria genomic sequences. AAA81453 to AAA82414 represent specifically claimed Neisseria meningitidis genomic DNA sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA sequences and their corresponding proteins; AAA81254 to AAA81259 and AAA81304 to AAA81321 representing PCR primers used in the isolation of Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF sequences, which are all used in the exemplification of the present invention. The nucleic acid sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to Neisserial bacteria. For example, some of the identified proteins could be components of vaccines against Meningococcus B, against all serotypes; and/or against all pathogenic Neisseriae. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious Meningococcus B vaccines have failed mainly due to antigen tolerance. Multivalent vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may		

CC provide an opportunity to identify secreted or surface exposed proteins  
CC that may be presumed targets for the immune system and which are not  
CC antigenically variable or at least more conserved than other more  
CC variable regions

XX Sequence 320 AA;

Query Match 99.3%; Score 1634; DB 3; Length 320;  
Best Local Similarity 99.1%; Pred. No. 1,4e-150;  
Matches 317; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRARLLPILFVSFVILSACGTLTGIPSHGGGKRRFAVEQELVAASARAAYKMDLQALHGR 60  
DB 1 MQARLLPILFVSFVILSACGTLTGIPSHGGGKRRFAVEQELVAASARAAYKMDLQALHGR 60  
QY 61 KVALYIATMGDQSGSLTGGRYSIDLIRGEYINSPARTDYTPRYETTAETTSGLTG 120  
DB 61 KVALYIATMGDQSGSLTGGRYSIDLIRGEYINSPARTDYTPRYETTAETTSGLTG 120  
QY 121 LTTSLSTLNAPALSRTOGSGSRSLGLNIGMGDYNETLTTNPRDTAFLSHLVOTVF 180  
DB 121 LTTSLSTLNAPALSRTOGSGSRSLGLNIGMGDYNETLTTNPRDTAFLSHLVOTVF 180  
QY 181 FLRGIDVVS PANADTVF INIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
DB 181 FLRGIDVVS PANADTVF INIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
QY 241 IKPKTNFAEAYKENYALMMGPYKSKGIRKPTGELMVDPSDIRPYGHTGNSAPVEADN 300  
DB 241 IKPKTNFAEAYKENYALMMGPYKSKGIRKPTGELMVDPSDIRPYGHTGNSAPVEADN 300  
QY 301 SHEGYGSDAVROHROGQP 320  
DB 301 SHEGYGSDAVROHROGQP 320

RESULT 12

AAB58586 ID AAB58586 standard; protein; 320 AA.

XX AAB58586;

XX 13-MAR-2001 (first entry)

XX N. meningitidis amino acid sequence m406.pap SEQ ID NO:75.

XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial.

XX Neisseria meningitidis.

OS WO200066791-A1.

XX 09-NOV-2000.

XX 08-MAR-2000; 2000WO-US005928.

XX 30-APR-1999; 99US-0132068P.

XX 08-OCT-1999; 99WO-US023573.

XX 28-FEB-2000; 2000GB-00004655.

XX (CHIR) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

PI Pizsa M, Hickey E, Peterson J, Tettelin H, Venter JC,  
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,  
PI Rappelli R, Frazer CM, Grandi G;

XX WPI; 2000-647603/62.

XX N-PSDB; AAF21581.

PT Neisseria meningitidis B full length genome sequence and open reading  
frames are used to detect, treat and prevent Neisserial infections.

XX Example 1; Page 108; 692pp; English.

CC The present invention describes the full length genome of Neisseria  
CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613  
CC represent fragments of the NMB genomic sequence, as the sequence was too  
CC long to go in a record on its own it was split into 8 sequences which  
CC overlap each other at the beginning and end of each sequence by 49980 bp  
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of  
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of  
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins  
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR  
CC primers which are used in the exemplification of the present invention.  
CC The NMB genome and fragments from it have antibacterial activity, and can  
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
CC and/or antibodies which binds to the proteins can be used in compositions  
CC for treating or preventing infection due to Neisserial bacteria or as a  
CC diagnostic reagent for detecting the presence of Neisserial bacteria or  
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,  
CC computer storage medium or computer databases can be used in a search to  
CC identify open reading frames (ORFs) or coding sequences within the NMB  
CC genome. The DNA sequences provide further opportunities to find antigenic  
CC or immunogenic proteins which are more effective in vaccines

XX Sequence 320 AA;

Query Match 99.3%; Score 1634; DB 3; Length 320;  
Best Local Similarity 99.1%; Pred. No. 1,4e-150;  
Matches 317; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRARLLPILFVSFVILSACGTLTGIPSHGGGKRRFAVEQELVAASARAAYKMDLQALHGR 60  
DB 1 MQARLLPILFVSFVILSACGTLTGIPSHGGGKRRFAVEQELVAASARAAYKMDLQALHGR 60  
QY 61 KVALYIATMGDQSGSLTGGRYSIDLIRGEYINSPARTDYTPRYETTAETTSGLTG 120  
DB 61 KVALYIATMGDQSGSLTGGRYSIDLIRGEYINSPARTDYTPRYETTAETTSGLTG 120  
QY 121 LTTSLSTLNAPALSRTOGSGSRSLGLNIGMGDYNETLTTNPRDTAFLSHLVOTVF 180  
DB 121 LTTSLSTLNAPALSRTOGSGSRSLGLNIGMGDYNETLTTNPRDTAFLSHLVOTVF 180  
QY 181 FLRGIDVVS PANADTVF INIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
DB 181 FLRGIDVVS PANADTVF INIDVFGTIRNRTMHLVNAETLKAQTKLEYFAVDRTNKKL 240  
QY 241 IKPKTNFAEAYKENYALMMGPYKSKGIRKPTGELMVDPSDIRPYGHTGNSAPVEADN 300  
DB 241 IKPKTNFAEAYKENYALMMGPYKSKGIRKPTGELMVDPSDIRPYGHTGNSAPVEADN 300  
QY 301 SHEGYGSDAVROHROGQP 320  
DB 301 SHEGYGSDAVROHROGQP 320

RESULT 13

AAY38539 ID AAY38539 standard; protein; 320 AA.

XX AAY38539;

XX 08-OCT-1999 (first entry)

XX Neisseria meningitidis strain A antigen encoded by ORF15.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW treatment; Neisseria infection; meningitis; septicæmia; gonorrhoea.

XX Neisseria meningitidis.

XX WO9924578-A2.

XX 20-MAY-1999.

```

XX 09-OCT-1998; 98WO-1B001665.
PR 06-NOV-1997; 97GB-00023516.
XX 14-NOV-1997; 97GB-00024190.
PR 18-NOV-1997; 97GB-00024386.
XX 27-NOV-1997; 97GB-00025158.
PR 10-DEC-1997; 97GB-00026147.
XX 14-JAN-1998; 98GB-00000759.
PR 01-SEP-1998; 98GB-00019016.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
PI WPI; 1999-327407/27.
XX N-PSDB; AA212007.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection.
XX
XX Claim 4; Page 98-99; 524pp; English.
XX
XX Amino acid sequences AAY38499-Y38944 represent Neisseria meningitidis and
CC N. gonorrhoeae antigenic proteins. They are encoded by open reading
CC frames (ORFs) AA211972-212358. The antigenic proteins, their fragments,
CC their nucleic acids and antibodies are used for diagnosis, prevention (as
CC vaccines) or treatment of Neisseria infections, such as meningitis,
CC septicemia and gonorrhea. Both organisms are closely related. Fragments
CC of the nucleic acids are useful as hybridisation probes and antisense
CC reagents
XX
XX SQ Sequence 320 AA;
XX
XX Query Match 98.7%; Score 1624; DB 2; Length 320;
XX Best Local Similarity 98.4%; Pred. No. 1.3e-149;
XX Matches 315; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MRALLILPLFSVFLSACGTLTGIPSHGGGRFAVEOELVAASARAAYKMDLQALHGR 60
DB 1 MOARLLILPLFSVFLSACGTLTGIPSHGGGRFAVEOELVAASARAAYKMDLQALHGR 60
QY 61 KVALYIATMGDGGSGSLTGRYSIDALRGEYINSPAVRTDYTPRYETTAETTSGLTGG 120
DB 61 KVALYIATMGDGGSGSLTGRYSIDALRGEYINSPAVRTDYTPRYETTAETTSGLTGG 120
QY 121 LTTSLSTLNAPALSRTOGDSGSRSSSLGINTGMDYRNELTTNPRDTAFSLHVOYVF 180
DB 121 LTTSLSTLNAPALSRTOGDSGSRSSSLGINTGMDYRNELTTNPRDTAFSLHVOYVF 180
QY 121 LTTSLSTLNAPALSRTOGDSGSRSSSLGINTGMDYRNELTTNPRDTAFSLHVOYVF 180
DB 121 LTTSLSTLNAPALSRTOGDSGSRSSSLGINTGMDYRNELTTNPRDTAFSLHVOYVF 180
QY 181 FLRGIDVVSPPANAADTDVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
DB 181 FLRGIDVVSPPANAADTDVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
QY 241 IKPKTNAPFAAYKENYALMGMGPYKVSKGIKPTEGLMVDPSDIQPYGNHNSAPSVEADN 300
DB 241 IKPKTNAPFAAYKENYALMGMGPYKVSKGIKPTEGLMVDPSDIQPYGNHNSAPSVEADN 300
QY 301 SHEGYGSDAEAVRHRQGP 320
DB 301 SHEGYGSDAEAVRHRQGP 320
XX
XX RESULT 14
XX AAY74928
XX ID AAY74928 standard; protein; 320 AA.
XX AC AAY74928;
XX
XX 21-MAR-2000 (first entry)
XX DT
XX DE Neisseria meningitidis ORF 406 protein sequence SEQ ID NO:1330.
XX

```

```

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX antibacterial; gene therapy.
XX
XX Neisseria meningitidis.
XX
XX WO957280-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US009346.
XX
XX 01-MAY-1998; 98US-0083758P.
PR 31-JUL-1998; 98US-0094869P.
XX 02-SEP-1998; 98US-0098994P.
PR 02-SEP-1998; 98US-0099062P.
XX 09-OCT-1998; 98US-0103749P.
PR 09-OCT-1998; 98US-0103794P.
XX 09-OCT-1998; 98US-0103796P.
PR 25-FEB-1999; 99US-0121528P.
XX
XX (CHIR-) CHIRON CORP.
XX (GENO-) INST GENOMIC RES.
XX
XX Frazer C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scarlato V, Scarselli M;
XX Teteltn H, Venter JC;
XX
XX WPI; 2000-062150/05.
XX N-PSDB; AA253690.
XX
XX Novel Neisseria polypeptides predicted to be useful antigens for
PT vaccines and diagnostics.
XX
XX Claim 2; Page 730; 1453pp; English.
XX
XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254773 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of the
CC invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the manufacture
CC of medicaments for treating or preventing infection due to Neisseria
CC bacteria (e.g. meningitis and septicemia), to detect the presence of
CC Neisseria bacteria, or to raise antibodies. They may also be used to
CC screen for agonists or antagonists, which may themselves have use as
CC antibacterial agents. The polynucleotides of the invention may also be
CC used in gene therapy protocols
XX
XX SQ Sequence 320 AA;
XX
XX Query Match 98.7%; Score 1624; DB 3; Length 320;
XX Best Local Similarity 98.4%; Pred. No. 1.3e-149;
XX Matches 315; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MRALLILPLFSVFLSACGTLTGIPSHGGGRFAVEOELVAASARAAYKMDLQALHGR 60
DB 1 MOARLLILPLFSVFLSACGTLTGIPSHGGGRFAVEOELVAASARAAYKMDLQALHGR 60
QY 61 KVALYIATMGDGGSGSLTGRYSIDALRGEYINSPAVRTDYTPRYETTAETTSGLTGG 120
DB 61 KVALYIATMGDGGSGSLTGRYSIDALRGEYINSPAVRTDYTPRYETTAETTSGLTGG 120
QY 121 LTTSLSTLNAPALSRTOGDSGSRSSSLGINTGMDYRNELTTNPRDTAFSLHVOYVF 180
DB 121 LTTSLSTLNAPALSRTOGDSGSRSSSLGINTGMDYRNELTTNPRDTAFSLHVOYVF 180
QY 181 FLRGIDVVSPPANAADTDVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
DB 181 FLRGIDVVSPPANAADTDVFINIDVFGTIRNRTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
QY 241 IKPKTNAPFAAYKENYALMGMGPYKVSKGIKPTEGLMVDPSDIRPYGNHNSAPSVEADN 300
DB 241 IKPKTNAPFAAYKENYALMGMGPYKVSKGIKPTEGLMVDPSDIRPYGNHNSAPSVEADN 300
XX

```

```

Db      241  |||
      241  IKPTNAFEAAAYKENVALMMGPYKVSKGIKPTGLMVDPSDIQPYGNHMGNSAPSEVADN 300
QY      301  |||
      301  SHEGYGSDAAYRHRGQOP 320
Db      301  |||
      301  SHEGYGSDAAYRHRGQOP 320

RESULT 15
AAV75813
ID      AAV75813 standard; protein; 320 AA.
XX
AC      AAV75813;
DT      21-MAR-2000 (first entry)
XX
DE      Neisseria meningitidis ORF 406 protein sequence SEQ ID NO:3114.
XX
KW      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW      antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW      antibacterial; gene therapy.
XX
OS      Neisseria meningitidis.
XX
PN      W09957280-A2.
XX
PD      11-NOV-1999.
XX
PF      30-APR-1999; 99WO-US009346.
XX
PR      01-MAY-1998; 98US-0083758P.
PR      31-JUL-1998; 98US-0094869P.
PR      02-SEP-1998; 98US-0098994P.
PR      02-SEP-1998; 98US-0099062P.
PR      09-OCT-1998; 98US-0103749P.
PR      09-OCT-1998; 98US-0103794P.
PR      09-OCT-1998; 98US-0103796P.
PR      25-FEB-1999; 99US-0121528P.
XX
PA      (CHIR ) CHIRON CORP.
PA      (GENO-) INST GENOMIC RES.
XX
PI      Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI      Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI      Tettelin H, Venter JC;
XX
DR      WPI; 2000-062150/05.
DR      N-PSDB; AAZ54615.
XX
PT      Novel Neisserial polypeptides predicted to be useful antigens for
PT      vaccines and diagnostics.
XX
PS      Example 1, Page 110-111; 1453pp; English.
XX
CC      AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAV74253 to AAV75941
CC      represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC      and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC      PCR primers used in the exemplification of the present invention. The
CC      polypeptides, the polynucleotides, antibodies and compositions of the
CC      invention can be used as vaccines, as diagnostic reagents, and as
CC      immunogenic compositions. The polypeptides can be used in the manufacture
CC      of medicaments for treating or preventing infection due to Neisserial
CC      bacteria (e.g. meningitis and septicemia), to detect the presence of
CC      Neisseria bacteria, or to raise antibodies. They may also be used to
CC      screen for agonists or antagonists, which may themselves have use as
CC      antibacterial agents. The polynucleotides of the invention may also be
CC      used in gene therapy protocols
XX
SO      Sequence 320 AA;

```

Query Match 98.7%; Score 1624; DB 3; Length 320;  
 Best Local Similarity 98.4%; Pred. No. 1.3e-149;  
 Matches 315; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

QY      1  MRARLLIPILFSVFILSACGLTGTGIPSHGGGKRFAVEQELVAASARAAYVXDMDLQALHGR 60
      1  :|||||
Db      1  MQARLLIPILFSVFILSACGLTGTGIPSHGGGKRFAVEQELVAASARAAYVXDMDLQALHGR 60
QY      61  KVALYIATMGDQSGSLTGGRYSIDALIIRGEYINSPAVRDYIYPRYETTAEITSSGGLTG 120
      61  |||
Db      61  KVALYIATMGDQSGSLTGGRYSIDALIIRGEYINSPAVRDYIYPRYETTAEITSSGGLTG 120
QY      121  LTTSLSTLNAPALSRTQSDGSSKSLGLNIGMGDYRNETLTNPDPDTAFSLHVOYTF 180
      121  |||
Db      121  LTTSLSTLNAPALSRTQSDGSSKSLGLNIGMGDYRNETLTNPDPDTAFSLHVOYTF 180
QY      181  FLRGIDVSPAMADTDVFINIDVFGTIRNRTKMLVNAETLKAQTKLEYFAVDRTNKKL 240
      181  |||
Db      181  FLRGIDVSPAMADTDVFINIDVFGTIRNRTKMLVNAETLKAQTKLEYFAVDRTNKKL 240
QY      241  IKPTNAFEAAAYKENVALMMGPYKVSKGIKPTGLMVDPSDIRPYGNHMGNSAPSEVADN 300
      241  |||
Db      241  IKPTNAFEAAAYKENVALMMGPYKVSKGIKPTGLMVDPSDIRPYGNHMGNSAPSEVADN 300
QY      301  SHEGYGSDAAYRHRGQOP 320
      301  |||
Db      301  SHEGYGSDAAYRHRGQOP 320

```

Search completed: August 18, 2005, 22:48:18  
 Job time : 131 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2005, 22:48:29 ; Search time 57 Seconds

(without alignments)  
540.164 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645  
Sequence: 1 MRARLLPILFVSFVILSACG.....SHEGYSDEAVRHRQGP 320

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1624	98.7	320	2	adhesin NMA0325 (i
2	1029	62.6	313	2	adhesin mafa - Nei
3	1026	62.4	313	2	adhesin mafa2 NMA2
4	104	6.3	863	2	Probable retroelem
5	102.5	6.2	756	2	chitinase B homolo
6	102	6.2	1004	2	surface layer prot
7	100	6.1	3147	2	hypothetical prote
8	99.5	6.0	756	2	chitinase B homolo
9	99.5	6.0	874	2	hypothetical prote
10	99	6.0	917	2	surface-layer prot
11	97.5	5.9	674	2	flagellar hook-ase
12	97	5.9	614	2	vitamin B12 recept
13	96.5	5.9	974	2	probable villin (i
14	96.5	5.9	1426	2	hypothetical prote
15	96	5.8	1343	2	hypothetical prote
16	95.5	5.8	802	2	xylanase, beta(1,3
17	95	5.8	685	2	probable flagellar
18	94.5	5.7	386	2	probable membrane
19	94	5.7	1205	2	hypothetical prote
20	93	5.7	451	2	uncharacterized co
21	93	5.7	456	2	hypothetical prote
22	93	5.7	474	2	flagellin - Legion
23	93	5.7	1042	2	sporid-like domain
24	93	5.7	1325	2	Ydek protein - Bac
25	93	5.7	1237	2	dextranase - Stre
26	93	5.7	1551	2	gramicidin S biosy
27	93	5.7	2383	2	probable membrane
28	92.5	5.6	747	2	cellulase (EC 3.2.
29	92	5.6	1037	2	plus-associated p

30	92	5.6	1051	2	S27002	phospholipase C (E
31	92	5.6	1234	2	S52099	phospholipase C be
32	92	5.6	1234	2	T38994	phospholipase C-be
33	92	5.6	1589	2	T42233	submaxillary mucin
34	92	5.6	3282	2	B62750	hemagglutinin-like
35	91.5	5.6	239	2	D61218	conserved hypothet
36	91.5	5.6	509	2	JC5651	N-acetylglucosamin
37	91.5	5.6	523	2	F82254	PTS system, N-acet
38	91.5	5.6	1029	2	S56229	probable membrane
39	91	5.5	5627	2	C83339	hypothetical prote
40	90.5	5.5	411	2	H75376	hypothetical prote
41	90.5	5.5	476	2	A63235	alkaline phosphata
42	90.5	5.5	799	2	S75351	hypothetical prote
43	90.5	5.5	807	2	AC2608	phenylalanyl-tRNA
44	90.5	5.5	807	2	B97390	phenylalanyl-tRNA
45	90	5.5	269	2	B64134	peptide transport

#### ALIGNMENTS

##### RESULT 1

adhesin NMA0325 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C:Accession: C82028  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holtroyd, S.; Jogle, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: AB1775; MUID:20222556; PMID:10761919  
A:Accession: C82028  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-320 <PAR>  
A:Cross-references: UNIPROT:O9JWK7; GB:AL162752; GB:AL157959; NID:97378778; PIDN:CAB8363  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: mafa; NMA0325  
C:Superfamily: Neisseria adhesin mafa

Query Match	Query	Match	Score	DB 2	Length	320
Best Local Similarity	98.4%	Pred. No. 1.5e-122				
Matches 315; Conservative 4; Mismatches 1; Indels 0; Gaps 0;						
QY	1	MRARLLPILFVSFVILSACGTLTGIPSHGGRFAVEQELVAASRAA VKMDLQALHGR	60			
DB	1	MOARLLPILFVSFVILSACGTLTGIPSHGGRFAVEQELVAASRAA VKMDLQALHGR	60			
QY	61	KVALYIATMGDSGSLTGRYSIDALRGYINSPARTYTYRYETLTTSGGLTG	120			
DB	61	KVALYIATMGDSGSLTGRYSIDALRGYINSPARTYTYRYETLTTSGGLTG	120			
QY	121	LTTSLTNAPALSRTOGSGSRSLGINTGMDYANETLTTPRDTAFSLHVQYVF	180			
DB	121	LTTSLTNAPALSRTOGSGSRSLGINTGMDYANETLTTPRDTAFSLHVQYVF	180			
QY	181	FLRGIDVSPANADTVFINIDVFTINRTEMLYNAETLKAQYKLEYPFADRNKLL	240			
DB	181	FLRGIDVSPANADTVFINIDVFTINRTEMLYNAETLKAQYKLEYPFADRNKLL	240			
QY	241	IKPKTNAFEALYKENVALMGYPYKSKIKPREGIMVDFSDIRPYGNHTGNSAPVEADN	300			
DB	241	IKPKTNAFEALYKENVALMGYPYKSKIKPREGIMVDFSDIRPYGNHTGNSAPVEADN	300			
QY	301	SHEGYSDEAVRHRQGP 320				
DB	301	SHEGYSDEAVRHRQGP 320				
RESULT 2						
	GB1173					

Query Match	62.4%;	Score 1026;	DB 2;	Length 313;
Best Local Similarity	62.4%;	Pred. No. 1.2e-74;		

chitinase B homolog lin0153 [imported] - *Listeria innocua* (strain C11p11262)  
C:Species: *Listeria innocua*

C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
 C/Accession: AB1452  
 R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fahl, H.D.; Jones, L.M.; Kars, U.  
 Science 294, 849-852, 2001  
 A:Author: Kref, J.; Kuhn, M.; Kunat, F.; Kurapat, G.; Madueno, E.; Maitouram, A.; Meok, C.; Schlueter, T.; Smoe, N.; Tietze, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.  
 A>Title: Comparative genomics of *Listeria species*.  
 A/Reference number: AB1077; MUID:21537279; PMID:11679669  
 A/Accession: AB1452  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-756 <G>  
 A/Cross-references: UNIPROT:Q92FF3; GB:AL592022; PIDN:CA95386.1; PID:g16412572; GSPDB:G  
 A/Experimental source: strain C1j11262  
 C/Genetics:  
 A:Gene: 11n0153  
 C/Superfamily: chitinase with C-terminal chitin-binding domain

Query Match 6.2%; Score 102.5; DB 2; Length 756;  
 Best Local Similarity 22.7%; Pred. No. 3.9;  
 Matches 73; Conservative 38; Mismatches 110; Indels 101; Gaps 16;

QY 18 ACGTGTGIPSHGCGKRAVEQELVAASAAVAKMDLQALHGRKVALYATMGDGGSL 77  
 DB 321 AAGDTALPG-----LFOAAEKTKNDGSLTYGANNENPKT---GDGR 363  
 QY 78 TGGRY-----SIDAL-----IRGEYINPAVTDYTPRYETTAETTSGLTG 120  
 DB 364 AGGVAAVYSIDLKAKPTLKEVMDTAKAPLYLKEKECEFTY----- 407  
 QY 121 LTTSLSTNAPALSTQDGGSSRSLGINIGMDY---RNEULLTPNPDATFLSHLVQ 177  
 DB 408 -----DWTBSIGYKAYQVQKNNLGMISMQSQKTTSTTRDE--LTAKK 452  
 QY 178 TTFVLRGIDVVSPPAADVDVFINIDVFGIRNRTEMLYNATLKAQKLEFAVDNRNK 237  
 DB 453 TGLF--GTSAL-PQNAIT--YANLVVAVKPKYSSNGVYEITITNNEK----ADETNE 502  
 QY 238 KLLIKPTNAFEAA-----YKENVALMGPYKSKIKPTEG--LMVDFSDI----- 282  
 DB 503 --VLKSTELSPFTVTLKPKYIPVKAGETLTAGDYKA--GTVTSSGNVTVDLASVYDAAQ 558  
 QY 283 ---RPGNHTGNSAPSYEADN 300  
 DB 559 IPOGASYTFRLKSSASVDVNN 580

## RESULT 6

T00046  
 surface layer protein - *Serratia marcescens*  
 C/Species: *Serratia marcescens*  
 C/Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004  
 C/Accession: T00046  
 R/Kawai, E.; Akatsuka, H.; Idei, A.; Shibata, T.; Omori, K.  
 Mol. Microbiol. 27, 941-952, 1998  
 A>Title: *Serratia marcescens* S-layer protein is secreted extracellularly via an ATP-bind  
 A/Reference number: Z14078; MUID:98194707; PMID:9535084  
 A/Accession: T00046  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1004 <K>  
 A/Cross-references: UNIPROT:Q54455; EMBL:AB007125; NID:92988355; PIDN:BAA25306.1; PID:92  
 A/Experimental source: strain str41; clone pSLA400  
 C/Genetics:  
 A:Gene: slsA

Query Match 6.2%; Score 102; DB 2; Length 1004;  
 Best Local Similarity 20.4%; Pred. No. 6.3;  
 Matches 67; Conservative 49; Mismatches 107; Indels 106; Gaps 16;  
 QY 56 ALHGKVALYATMGDGG---SLTGGYSLDALIRGEYINSPAVTD----- 101

DB 5 AQTNNVAMTYAIIQKANKVFPDLOAONIYGLQTTSDPNTVTSTDGALYAKGSDS 64  
 QY 102 -----YT-----YPRYETTAETTSGLTGLTTSSTLN---APALSRTQS--- 138  
 DB 65 DILTTIRKVVAVASSTVLSGVLNGTTLAQNIALSLTINLVTVNGPSTLLSQQSFOQ 124  
 QY 139 -----DGGSSRSL-GLNIGG---MGDRNETL-----TTNPD----- 168  
 DB 125 ILNTALVADGAMTGLGGTYKGLAVYNGVPTDAILQSIGTGINGGTAETDAATALLQ 184  
 QY 169 -----TAFSLVQTF---FLRGIDVSPNADTDVFINIDVFGIRNRTEMLYNAA 219  
 DB 165 LRPGLVTAATNPFVKSIRKAGLR-----DPTTELNGVIELSSGASRAQYIVAVY--D 237  
 QY 220 TLK-----AQTU-----EYFAVDNRNKLLIKPTNAFEAAKYNALMWG 261  
 DB 238 TLKTTTSGSDLAQKVLVSSPSYNDGYLPSMAKQELASLYLVFD---RQIDALGIS 294  
 QY 262 PY--KVSKGIRTEGLMVPFSDIRPYGNH 288  
 DB 295 SYTRQLERGANVYDGIKEFLVSRFGIH 323

## RESULT 7

T18674  
 hypothetical protein T04F3.1 - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T18674; T24464  
 R.White, S.

submitted to the EMBL Data Library, June 1996  
 A/Reference number: Z19004  
 A/Accession: T18674  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3147 <W>  
 A/Cross-references: UNIPROT:Q17464; EMBL:Z74026; PIDN:CAA98419.1; GSPDB:GN00023; CESP:T0  
 A/Experimental source: clone B0240  
 R.Kershaw, J.  
 submitted to the EMBL Data Library, May 1996  
 A/Reference number: Z19894  
 A/Accession: T24464  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3147 <W12>  
 A/Cross-references: EMBL:Z72513; PIDN:CAA96672.1; GSPDB:GN00023; CESP:T04F3.1  
 A/Experimental source: clone T04F3  
 C/Genetics:  
 A:Gene: CESP:T04F3.1  
 A:Map position: 5  
 A:Insertions: 338/3; 417/3; 1957/2; 2358/3; 2400/1; 2529/2; 2585/1; 2700/3; 2744/1; 2877/3;

Query Match 6.1%; Score 100; DB 2; Length 3147;  
 Best Local Similarity 22.7%; Pred. No. 44;  
 Matches 58; Conservative 40; Mismatches 94; Indels 64; Gaps 11;

QY 68 TMDGGSGSLTGG---RYS-----IDALIRGEYINSPA-----VRTDYTPRYE 108  
 DB 2477 TIGSEDSGMSGGGRGRRSFRITDQVIDEAFQGFDSQSTSTNPKPVRTEHYDDY 2536  
 QY 109 TTA-----ERTSGTLTSLTSLT-----NAPALSRTQS-----DGGSSRSL 147  
 DB 2537 ITSLQOEDLDATDEVDENLDVSTFVDLILGKSMDEAFLSSTKSLREHTDTSIDRKKS 2596  
 QY 148 GLNIGMGMDYNETLT-----TNPRDTAFSLVQTFELRGIDVSPNADTD 196  
 DB 2597 GEKVHAY--YNNRDTSDTSIDKSKPEVITDELTSELDEINKVY-----VEPSVSKSD 2648  
 QY 197 VFINDVFGIRNRTEMLYNAAETLKAQTKLEYFAVDRTNKKLLIK--PKTNAAFEAAVKE 254  
 DB 2649 SSANIK--ASQNSTKTPCDEBLELEIKSEYFLIKGSYSLILPKSPDLKQMLKQLEQ 2705



## C:Genetics:

A:Gene: alpha

F:1-40/Domain: signal sequence #status predicted <SIC>  
F:25-79/Domain: S-layer repeat homology <SLR1>  
F:41-917/Product: S-layer protein #status predicted <MAT>

Query Match 6.0%; Score 99; DB 2; Length 917;

Best Local Similarity 22.9%; Pred. No. 9.6; Mismatches 118; Indels 122; Gaps 22;  
Matches 84; Conservative 43; Mismatches 118; Indels 122; Gaps 22;

QY 21 TLTGIPSHGGKR--FAVEQELVAASARAAYKMDLQALHGRVALYATMGDGGSLT 78  
DB 444 TLTVGYGSGKANGNGPFTYR-----GARLALSPVEGIALGGS----FVGEGLDANQGTTS 494  
QY 79 GG-----RYSIDALIR-----GEYINSPA-----VTTYTPRY----- 107  
DB 495 ASFPATTVYGVDAVSKGPGVGLAGEYFNSDAAPNANGYVADVALGSIYAGNYRNIG 554  
QY 108 -----BTPTETTSGLTGLTTSLTSLN-APALSRTQSDGSGRSSLG-INIGGM 154  
DB 555 AGVTGANNLSDPATSTLDGGGNGVDSGNGVINGAPFPNSNRGFGVSAAGLPITVKGY 614  
QY 155 GD-----YNNETLTN-----PRDTAFLSHLVQTVFPLRGIDVSPANA 193  
DB 615 YDSSTVLANETITNSYGFANVSANNQVAVGGQDLAFGCG-----FTLSGFYRIAQLNG 668  
QY 194 DTDFINIDVFGTINRRTBMLYNAETLKAQTKLEYFAVDRTNKKLLIKETNAFEAYK 253  
DB 669 STTRYI-----LTERK---PAEAVYASEYGAKLHMDGASKDALV-PKLN-FTAAVT 713  
QY 254 ENY-----ALMMGPYKSKGIKPTREGLVWDFSDIRPYGN-HTGN-SAPSVLEAD 299  
DB 714 QKYNATGSGFTTQDLAIV-GSYELALG-PL-----TLKPMGRYHTQDMAASTSD 762  
QY 300 NSHEGYG 306  
DB 763 YTVKYG 769

## RESULT 11

H64613 flagellar hook-associated protein 2 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004

C:Accession: H64613

R:Tomd, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen-

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C.

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: H64613

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-674 &lt;TOM&gt;

A:Cross-references: UNIPROT:P96786; GB:AE000587; GB:AE000511; NID:G2313869; PIDD:AAD0780

Query Match 5.9%; Score 97.5; DB 2; Length 674;

Best Local Similarity 20.7%; Pred. No. 8.3; Mismatches 120; Indels 75; Gaps 13;  
Matches 62; Conservative 43; Mismatches 120; Indels 75; Gaps 13;

QY 37 EQELVAASARAAYKMDQ--ALHGRKALYATMGDGGSLTGKYSIDALIRGEYI 93  
DB 317 ESDLLKSS--RTIKGKLEGVSLNGQKDLKSALTRESNTSENT-----DAIIQA--I 366  
QY 94 NSPAVRTYTYPRYETTAETTSGLTGLTTSLTSLNAPALSRQSDGSGRSSLGANIGG 153  
DB 367 NAKESLSAFKAKAGGLVINSKTKMLT-----IKEDALGKASLKDLGANNKM 413  
QY 154 MGDYNETLTNPRDTAFLSHLVQ---TVFPLRGIDVSPANADTVFINIDVFGTIRN 209  
DB 414 VOSYBAS-----QNTLFMSKNLQKASDSAFYNGVSIIRPTNEVDVNSGVNI--TLLEQ 465

QY 210 RTEMLYNAETLKAQTKLEYFAVDRTNKKLLIKETNAFEAYKE-----HYALMM 260  
DB 466 TTE-----DNKPAIISVSXDNDPAII--DSLTEFYKAYNLIPKLEDTETRYDADT 512  
QY 261 GPYKVSXGKIKPTREGLVWDFSDIRPYGNHTGNSAPSEVADNSHEGYGS-----DEA 311  
DB 513 KIAGIFNGVGDIRAIRSSLNVNVSYSVHTDNQVESLMK-----YGLSIDDGKWSLDBA 566

## RESULT 12

AB0935

vitamin B12 receptor protein [imported] - Salmonella enterica subsp. enterica serovar Typh

C:Species: Salmonella enterica subsp. enterica serovar Typh

A:Note: This species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AB0935

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davelee, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.,

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AB0935

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-614 &lt;PAP&gt;

A:Cross-references: GB:AL513382; PIDD:CAD09500.1; PID:q16504617; GSPDB:GN00176

C:Genetics: C:Superfamily: vitamin B12 receptor; conb-dependent receptor amino-terminal homology; tc

Query Match 5.9%; Score 97; DB 2; Length 614;

Best Local Similarity 20.0%; Pred. No. 8; Mismatches 56; Conservative 42; Mismatches 116; Indels 66; Gaps 8;

QY 74 SGLTGRYSIDALIRGEYINSPAVRTYTYPRYETTAETTSGLTGLTSLNAPAL 133  
DB 111 SGSADLSQFPVSLVGRIRYINGP-----RSALYGSADIGAIVNIITTDNERT 158  
QY 134 SRTQSDGSGSRSSSLGNTIGMGDYNETLTN-----NPRDTAFLSHLVQTVFPLRGIDV 187  
DB 159 ELTACMGNS-----YGNVLSLQOQUGENTRAT-----LIGVEYTKGFV 200  
QY 188 VS-----PANADTVFINIDVFGTIR-----NRTMHLNA----- 218  
DB 201 VAKGSGTGMQAPDRDGFSLKTLYGALERTFSDRWSGFVRGYGYDNRDYDAYSPGSPLI 260  
QY 219 ETLKQTKLEYFAVDRTNKKLLIKETNAFEAYKENYALMMGPYKVSXGKIKPTREGLVWD 278  
DB 261 DTRKLYSQ-SWDAGLRFNGERIQSOLVSYSKDYNDPHYGRYDTSATLDEMKNQYVQ 319  
QY 279 FSDIRPYGNHTGNSAPSEVADNSHEGYGSDPAVYRQHOG 318  
DB 320 WTNISVYVGHGNGVAGVDWQKQTTTGTGYVREGYDQRTG 359

## RESULT 13

T14076 probable villin [imported] - Arabidopsis thaliana

N:Alternate names: protein P6G3.190; protein P9N1.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T14076; T08992; T5129

R:Beyan, M.; Kalicki, J.; Wohlmann, P.; Smith, A.; Bancroft, I.; Mewes, H.W.; Mayer, K.

submitted to the Protein Sequence Database, August 1999

A:Reference number: Z17873

A:Accession: T14076

A:Molecule type: DNA

A:Residues: 1-974 &lt;BEV&gt;

A:Cross-references: UNIPROT:O65570; EMBL:AL109796; GSPDB:GN00062; ATSP:F9N11.10  
A:Experimental source: cultivar Columbia; BAC clone F9N11



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2005, 22:29:44 ; Search time 129 Seconds  
(without alignments)  
1270.274 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645  
Sequence: 1 MRARLLIPILFVFLSACG.....SHEGYGSDAVRHRQGP 320

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1645	100.0	320	2	09X6P2
2	1624	98.7	320	2	09JWK7
3	1029	62.6	313	2	09J544
4	1026	62.4	313	2	09J5W0
5	122.5	7.4	511	2	08VP63
6	117.5	7.1	729	2	06NP01
7	117.5	7.1	1250	2	002434
8	117.5	7.1	1250	2	09VY88
9	117.5	7.1	1250	2	09VY88
10	114.5	7.0	627	2	08KG07
11	111.5	6.8	3552	2	08XSD6
12	109	6.6	874	2	06G428
13	107	6.5	684	2	06BCF8
14	105.5	6.4	731	2	06AVK7
15	104	6.3	863	2	09SKF9
16	102.5	6.2	756	2	09ZFF3
17	102.5	6.2	1240	2	07RS47
18	102.5	6.2	1265	2	08EV72
19	102	6.2	1004	2	054455
20	101	6.1	429	2	07PRY0
21	101	6.1	545	2	07OLB8
22	101	6.1	1326	2	08L898
23	101	6.1	3346	2	07WN54
24	100.5	6.1	496	2	06AN08
25	100.5	6.1	615	2	08J0W2
26	100.5	6.1	1096	2	08BKW6
27	100	6.1	1035	2	033372
28	100	6.1	3167	2	017464
29	99.5	6.0	756	2	08YAL3
30	99.5	6.0	874	2	09RRU3
31	99.5	6.0	1056	2	08A7L7

32	99.5	6.0	1500	2	09BKH1	09bkh1 drosophila
33	99	6.0	732	2	06ZB62	06z62 eynochocyt
34	99	6.0	794	2	07WZD3	07wzd3 salmonella
35	99	6.0	917	1	SLAP_THETH	P35830 thermus the
36	98.5	6.0	615	2	08J0Y2	08j0y2 cryptococcu
37	98.5	6.0	883	2	09S526	09s526 drosophila
38	98.5	6.0	1500	2	09V4C8	09v4c8 drosophila
39	98	6.0	612	2	05Z903	05z903 citrobacter
40	98	6.0	1175	2	08RHT3	08rht3 fusobacteri
41	98	6.0	3711	2	06FDW3	06fdw3 acinetobact
42	97.5	5.9	443	2	0890Y8	0890y8 clostridium
43	97.5	5.9	684	1	FLID_HELPPY	P96788 helicobacte
44	97.5	5.9	1507	2	06WIA0	06wia0 bdellovibri
45	97.5	5.9	2363	2	06ALH2	06alh2 desulfotale

## ALIGNMENTS

RESULT 1	ID	PRELIMINARY;	PRT;	320 AA.
09X6P2	09X6P2			
AC	09X6P2;			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Adhesin MaFA.			
GN	Name=mafa;			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;			
OC	Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=485;			
RN	(1)			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=MS11;			
RA	Eickertjaeger S., Meyer T.F., Fischer E., Maier J., Manning P.A.,			
RA	Rudel T., Scheuerpflug I., Schulz E., Schwan E.T.;			
RL	Submitted (Apr-1999) to the EMBL/Genbank/DDJ databases.			
DR	EMBL, AF142582; AAD31038.1; ..			
SQ	SEQUENCE 320 AA; 34779 MW; EE37DEC55E27DCE3 CRC64;			
Query Match	100.0%; Score 1645; DB 2; Length 320;			
Best Local Similarity	100.0%; Pred. No. 5.1e-125;			
Matches	320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MRARLLIPILFVFLSACGTLTGIPSHGGKRFVDEGLVAASARAAMKMDLALHGR	60	
DB	1	MRARLLIPILFVFLSACGTLTGIPSHGGKRFVDEGLVAASARAAMKMDLALHGR	60	
QY	61	KVALYIATMGOGSGSLTGRYSIDALRGELINSPARTDTYRYETTAETTSGLTGG	120	
DB	61	KVALYIATMGOGSGSLTGRYSIDALRGELINSPARTDTYRYETTAETTSGLTGG	120	
QY	121	LTTSLSTINAPALSTRQDSGSRSSLGINIGMDYENETLTTPRDTAFSLHVGTVF	180	
DB	121	LTTSLSTINAPALSTRQDSGSRSSLGINIGMDYENETLTTPRDTAFSLHVGTVF	180	
QY	181	FLRGIDVVS PANADTVFINIDVFETINRREMLYNAETLKAQTKLEYFAVDRTNKKLL	240	
DB	181	FLRGIDVVS PANADTVFINIDVFETINRREMLYNAETLKAQTKLEYFAVDRTNKKLL	240	
QY	241	IKPKTNAFEAYKENVALMGPKYKSKIKPREGIMVDPDRIKPNHTGNSAPVEADN	300	
DB	241	IKPKTNAFEAYKENVALMGPKYKSKIKPREGIMVDPDRIKPNHTGNSAPVEADN	300	
QY	301	SHEGYGSDAVRHRQGP 320		
DB	301	SHEGYGSDAVRHRQGP 320		
RESULT 2				
Q9JWK7	Q9JWK7	PRELIMINARY;	PRT;	320 AA.

AC Q9JMK7; (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Adhesin.  
 GN Name=mafa; OrderedLocustNames=NMA0325;  
 OS Neisseria meningitidis (serogroup A).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 NCBI\_TaxID=65699;  
 RX MEDLINE=22491 / Serogroup A / Serotype 4A;  
 RX PubMed=20222556; PubMed=10761919; DOI=10.1038/35006655;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,  
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,  
 Jørgensen K., Leather S., Moulé S., Mungall K.L., Quail M.A.,  
 Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 Whitehead S., Spratt B.G., Barrell B.G.;  
 RA "Complete DNA sequence of a serogroup A strain of Neisseria  
 meningitidis Z2491."  
 RT Nature 404:502-506(2000).  
 RL EMBL; AL162752; CAB83630.1; -.  
 DR PIR; C82028; C82028.  
 KW Complete proteome.  
 SQ SEQUENCE 320 AA; 34753 MW; C45EBBA8EECDE49 CRC64;

Query Match 98.7%; Score 1624; DB 2; Length 320;  
 Best Local Similarity 98.4%; Pred. No. 2,5e-123;  
 Matches 315; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRARLLILPILSVFLLSACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGR 60  
 DB 1 M0ARLLILPILSVFLLSACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGR 60  
 OY 61 KVALYIATMGDGGSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGLTJG 120  
 DB 61 KVALYIATMGDGGSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGLTJG 120  
 OY 121 LTTSTLNAPALSTQSDGSGRSLSGLNTGMDYNETLTTPRDTAFLSHVQTF 180  
 DB 121 LTTSTLNAPALSTQSDGSGRSLSGLNTGMDYNETLTTPRDTAFLSHVQTF 180  
 OY 181 FLRGIDVVSANADTVFINIDVFGTINRTTETMLYNAETLKAQTKLEYFVADRINKKL 240  
 DB 181 FLRGIDVVSANADTVFINIDVFGTINRTTETMLYNAETLKAQTKLEYFVADRINKKL 240  
 OY 241 IKPRTNAFEAAYKENYALMGPYKVKSGIKPTGIMVDFSDIRPYGNHTGNSAPSEADN 300  
 DB 241 IKPRTNAFEAAYKENYALMGPYKVKSGIKPTGIMVDFSDIRPYGNHTGNSAPSEADN 300  
 OY 301 SHEGYSDEAVRHHOGOP 320  
 DB 301 SHEGYSDEAVRHHOGOP 320  
 RESULT 3  
 O9JSMO PRELIMINARY; PRT; 313 AA.  
 AC Q9JSMO; (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE MafA protein.  
 GN OrderedLocustNames=NMB0375, NMB0652;  
 OS Neisseria meningitidis (serogroup B).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 NCBI\_TaxID=491;  
 RX MEDLINE=22491 / Serogroup A / Serotype 4A;  
 RX PubMed=20222556; PubMed=10761919; DOI=10.1038/35006655;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,  
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,  
 Jørgensen K., Leather S., Moulé S., Mungall K.L., Quail M.A.,  
 Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 Whitehead S., Spratt B.G., Barrell B.G.;

RX MEDLINE=20175755; PubMed=10710307; DOI=10.1126/science.287.5459.1809;  
 RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,  
 RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.C.,  
 RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,  
 RA Hickey E.K., Harte D.H., Salzberg S.L., White O., Fleischmann R.D.,  
 RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,  
 RA Citronne H., Clark E.B., Cotton M.D., Uterback T.R., Khouri H.M.,  
 RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Piza M.,  
 RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,  
 RA Venter J.C.;  
 RA "Complete genome sequence of Neisseria meningitidis serogroup B strain  
 MC58."  
 RT Science 287:1809-1815(2000).  
 RL EMBL; AE002393; AAF62309.1; -.  
 DR EMBL; AE002420; AAF41073.1; -.  
 DR PIR; G81173; G81173.  
 DR TIGR; NMB0375; -.  
 DR TIGR; NMB0652; -.  
 KW Complete proteome.  
 SQ SEQUENCE 313 AA; 34000 MW; 360676272DAD60C1 CRC64;

Query Match 62.6%; Score 1029; DB 2; Length 313;  
 Best Local Similarity 62.7%; Pred. No. 4.3e-75;  
 Matches 197; Conservative 49; Mismatches 62; Indels 6; Gaps 2;

OY 5 LLIPILSVFLLSACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGRKVAL 64  
 DB 6 LLIPILSVFLLSACGTLTGIPSHGGGRFAVEOELVAASARAANKMDLQALHGRKVAL 60  
 OY 65 YIATMGDGGSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGLTJG 124  
 DB 61 YIATMGDGGSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGLTJG 120  
 OY 125 LSTLNAPALSTQSDGSGRSLSGLNTGMDYNETLTTPRDTAFLSHVQTF 184  
 DB 121 LSTLNAPALSTQSDGSGRSLSGLNTGMDYNETLTTPRDTAFLSHVQTF 180  
 OY 185 IDVVSANADTVFINIDVFGTINRTTETMLYNAETLKAQTKLEYFVADRINKKL 244  
 DB 181 IEVPEPADTVFINIDVFGTINRTTETMLYNAETLKAQTKLEYFVADRINKKL 240  
 OY 245 TNAFEAAYKENYALMGPYKVKSGIKPTGIMVDFSDIRPYGNHTGNSAPSEADN 304  
 DB 241 TAAEYQYOEYALMTGPYKVKSKTVDRLMDFSDITPYGDTAQRDPKONGKKP 300  
 OY 305 YGYSDEAVRHHOG 318  
 DB 301 -DVGNEVIRRRKGG 313

RESULT 4  
 O9JSMO PRELIMINARY; PRT; 313 AA.  
 AC Q9JSMO; (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Adhesin MafA2.  
 GN Name=mafa2; OrderedLocustNames=NMA2112;  
 OS Neisseria meningitidis (serogroup A).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 NCBI\_TaxID=65699;  
 RX MEDLINE=22491 / Serogroup A / Serotype 4A;  
 RX PubMed=20222556; PubMed=10761919; DOI=10.1038/35006655;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,  
 Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,  
 Jørgensen K., Leather S., Moulé S., Mungall K.L., Quail M.A.,  
 Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 Whitehead S., Spratt B.G., Barrell B.G.;



RT "Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22481.".  
 RL Nature 404:502-506(2000).  
 DR EMBL; AL162758; CAB85326.1; -.  
 DR PIR; G81782; G81782.  
 KW Complete proteome.  
 SQ SEQUENCE 313 AA; 33929 MW; DFB76F2C537C5D7D CRC64;

Query Match 62.4%; Score 1026; DB 2; Length 313;  
 Best Local Similarity 62.4%; Pred. No. 7,5e-75;  
 Matches 196; Conservative 49; Mismatches 63; Indels 6; Gaps 2;

QY 5 LLIPLFVFLISACGTLTGIFSHGGRFAVEQLVAASARAAYKMDLALHGRKVAL 64  
 DB 6 LLIPLFVFLISACGTLTGIFSHGGRFAVEQLVAASARAAYKMDLALHGRKVAL 60  
 QY 65 YIATMGDGGSLTNGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTGLTTS 124  
 DB 61 YVSVNGDGGSLTNGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTGLTTS 120  
 QY 125 LSTLAPALSRQTSGSGSSRLGLNIGMGDYNRETLTTPRDTAFSLHVLQVYFPLRG 184  
 DB 121 TSLTAPALSRQTSGSGSSRLGLNIGMGDYNRETLTTPRDTAFSLHVLQVYFPLRG 180  
 QY 185 IDVSPANADTDVFINIDVFGTIRNRTMHLNNAETLKAQTKLEYFAVDRTMKLLKPK 244  
 DB 181 IEVPEYADTDVFINIDVFGTIRNRTMHLNNAETLKAQTKLEYFAVDRTMKLLKPK 240  
 QY 245 TNAPFAAYKENYALMMGPKYKSKGKIPTEGLMVFSDIRPYGNHTGNSAPSEADNSHG 304  
 DB 241 TAAVSGQOEOYALMMGPKYKSKGKIPTEGLMVFSDIRPYGNHTGNSAPSEADNSHG 300  
 QY 305 YGYSDEAVRQRHOG 318  
 DB 301 -DVGNVIRRRKGG 313

## RESULT 5

ID 08VP63 PRELIMINARY; PRT; 511 AA.  
 AC 08VP63;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Alkaline phosphatase.  
 GN Name:phoA;  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1772;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22778081; PubMed=12897018;  
 RX DOI=10.1128/JB.185.16.4983-4991.2003;  
 RA Kriakov J. I., Lee S. H., Jacobs W. R. Jr.;  
 RT "Identification of a regulated alkaline phosphatase, a cell surface-associated lipoprotein, in *Mycobacterium smegmatis*.";  
 RL J. Bacteriol. 185:4983-4991(2003).  
 CC -1 SIMILARITY: Belongs to the alkaline phosphatase family.  
 DR EMBL; AY065934; AAL50185.1; -.  
 DR HSBP; P06634; IAJA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR Pfam; PF00245; Alk\_phosphatase; 1.  
 DR PRINTS; PR00113; ALKPHPTASE.  
 DR SMART; SM00098; alkpc; 1.  
 DR PROSITE; PS00123; ALKALINE PHOSPHATASE; 1.  
 SQ SEQUENCE 511 AA; 53697 MW; FCB27A3A561C92C CRC64;

Query Match 7.4%; Score 122.5; DB 2; Length 511;  
 Best Local Similarity 21.7%; Pred. No. 0.26;  
 Matches 75; Conservative 41; Mismatches 105; Indels 125; Gaps 17;  
 QY 1 MRARLLILFSLFISACGTLTGIFSHGGRFAVEQLVAASARAAYKMDLALHGRKVAL 56

DB 7 LRATVVI-VAALALPLTCCSDTTTSSGDPVHRADDITYNGARRIDGQDTTERSA 65  
 QY 57 LHGRKVALYIATMGD-QGSGSLT-----GGRY-SIDAL-IRGEYINSPAVRTDYTY 104  
 DB 66 INGGKARVILLVGGMDSEITMARNYEKAGGSGFDLALPLSGQY-----TLYAL 118  
 QY 105 -----PRYETTAETTSGLTGLTSLTLNMP-----ALSRQSDGSG----- 142  
 DB 119 NKDGPNVTVTSASA--TGMVTGTYKYNALGIDIKGNPDKTLELAKOGFATGVT 175  
 QY 143 -----SRSSL-----GLNIGMGDYNRETLTTPRDTAFSL 173  
 DB 176 TSEIADATSAFLFHSIRDCYGVQTADCAADALENGSGSYEQTLTRP----- 228  
 QY 174 HLVTQVFFLRGIDVSPANADTDVFINIDVFGTIRNRTMHLNNAETLKAQTKLEYFAVD 233  
 DB 229 -----DLTLGGAEY-----FSQTNAGE-----YKQKTLQAQKRGFOIV 265  
 QY 234 RTNKKLLIKPKTNAPFAAYKENYAL-----MMGPKYKSKG 268  
 DB 266 RTASFL-----DAVTADQDNVYGLFADGNMPPVMSGPPAVRHG 305

## RESULT 6

ID 06NP01 PRELIMINARY; PRT; 729 AA.  
 AC 06NP01;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE HL04242p (Fragment).  
 GN Name:rdgB;  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkelley;  
 RA Stapleton M., Brokstein P., Hong L., Asgayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frisbe E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liso G.,  
 RA Miranda A., Mungall C.D., Nunoo J., Paclele J., Paragae V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiniker S.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BT011130; AAR82797.1; -.  
 DR GO; GO:0046872; P:metal ion binding; IEA.  
 DR InterPro; IPR004177; DDHD.  
 DR Pfam; PF02862; DDHD; 1.  
 FT NON TER 1  
 SQ SEQUENCE 729 AA; 79585 MW; 355F7B37ABE3A995 CRC64;

Query Match 7.1%; Score 117.5; DB 2; Length 729;  
 Best Local Similarity 23.7%; Pred. No. 1;  
 Matches 82; Conservative 39; Mismatches 106; Indels 119; Gaps 18;  
 QY 14 FILSACGTLTGIFSHGGR-----KRFVBOELVAASARA--AVKD 51  
 DB 394 FILRQIGKEBIGIFPGSDDKDNASFHRCQREKKIKRTSVKLKNVAANRAVDVIOE 453  
 QY 52 -----MDLALHGRKVALYIATMGDGGSGSLT-----GGR--YSID- 85  
 DB 454 GREORLNARFMYGGLMDITLHGEKVDVHI--MKDPPAGEWTFLESTVDKNGRISYSIPD 511  
 QY 86 --ALIRGEYINSPAVRTDYTY-----PRYETTAETTSGLTGLTSLTSL-FLNAPAL 133  
 DB 512 QVSLAGYGYPRKMYVRGDHSTVDCYMAVVPFLTECVPSIDG---SFLASNSYGRDPRK 568  
 QY 134 SRTQSDGSGSSSLG---LNIIGMGDYNRETL-----LITNP--RD 168

Db 569 RAGAVDVCRRHQEGLYLLIYITGRPDMQQRVSWLSQHNPHGLISFADGLSTDPGLGHK 628  
Qy 169 TAFSLHVQTVFPLRGIDVVS PANADTVFINIDVFGTIRNRTEMLHYNMETLKAQRTLE 228  
Db 629 TAYLNNLVON---HGSIITAYGSSKD---ISVYINVGKRTD-----Q 665  
Qy 229 YFAVDRTNKULLIKPTNAFEAYKENYALMGPYKYSKGIKPTG 274  
Db 666 IFIVGKVGKTL---QSNV---TVLSDGYAAHLGLAQVGSGRPAKG 705  
RESULT 7  
002434 PRELIMINARY; PRT; 1250 AA.  
AC 002434  
DT 01-JUN-1997 (TReMBLrel. 04, Created)  
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Retinal degeneration B protein.  
GN Name=rdgb;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R; TISSUE=Head;  
RX MEDLINE=91231170; PubMed=1903119;  
RA Vithelie T.S., Hyde D.R., O'Tousa J.E.,  
RT "Isolation and characterization of the Drosophila retinal degeneration  
B (rdgb) gene."  
RL Genetics 127:761-768 (1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R; TISSUE=Head;  
RX MEDLINE=98343889; PubMed=9680295;  
RA Rubboli F., Bulfone A., Bogni S., Marchitelli A., Zolli M.,  
RA Porsani G., Ballabio A., Banfi S.,  
RT "A mammalian homologue of the Drosophila retinal degeneration B gene:  
implications for the evolution of phototransduction mechanisms."  
RL Genes Funct. 1:205-213(1997).  
RN EMBL; Y08035; CAA69291.1; -.  
DR HSSP; P53810; 1KCM.  
DR FlyBase; FBgn0003218; rdgb.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0008525; F:phosphatidylcholine transporter activity; IDA.  
DR GO; GO:0008526; F:phosphatidylinositol transporter activity; IDA.  
DR GO; GO:0016059; P:deactivation of rhodopsin mediated signaling; IMP.  
DR GO; GO:0007608; P:perception of smell; IMP.  
DR GO; GO:0007602; P:phototransduction; TAS.  
DR GO; GO:0016056; P:rhodopsin mediated signaling; IMP.  
DR InterPro; IPR004177; DDHD.  
DR InterPro; IPR001666; P1\_transfer.  
DR Pfam; PF02862; DDHD; 1.  
DR Pfam; PF02121; IP\_trans; 1.  
DR PRINTS; PR00391; FITRANSFER.  
SQ SEQUENCE 1250 AA; 137777 MW; 3CC926ABBA0A6F28 CRC64;  
Query Match 7.1%; Score 117.5; DB 2; Length 1250;  
Best Local Similarity 23.7%; Pred. No. 2.1;  
Matches 82; Conservative 39; Mismatches 106; Indels 119; Gaps 18;  
Qy 14 FILSACCTGTCIPBHGGS-----KRFVDEIVAAASRA---AVKD 51  
Db 906 FILKQIKFESGIPVGSNDKDNASPHDGPREFWIKKRISSVKLVNAAANRADVIVQE 965  
Qy 52 -----MDQLALHGRKVALYIATMGDQSGSGLT-----GGR---YSID- 85  
Db 966 GREORLNAEFMYGSLDMITLHGEKVDVHI--MKRPPAGQWFFLSTVTDKNGRISYSIPD 1023  
Qy 86 --ALIRGEYINSPAVRTDYTY-----PRYETTAETTSGLTGLTSLIS-TLNAAPAL 133

Db 1024 QVSLGYGIYEPKAVVRDGHTSVDCYMAVBPPLRECVFISDG---SEFASMSVYGRDPKV 1080  
Qy 134 SRTOSSDSSGRSSLSG---LNIIGMGDYRNET-----LTTNP--RD 168  
Db 1081 RAGAVDVCRRHQEGLYLLIYITGRPDMQQRVSWLSQHNPHGLISFADGLSTDPGLGHK 1140  
Qy 169 TAFSLHVQTVFPLRGIDVVS PANADTVFINIDVFGTIRNRTEMLHYNMETLKAQRTLE 228  
Db 1141 TAYLNNLVON---HGSIITAYGSSKD---ISVYINVGKRTD-----Q 1177  
Qy 229 YFAVDRTNKULLIKPTNAFEAYKENYALMGPYKYSKGIKPTG 274  
Db 1178 IFIVGKVGKTL---QSNV---TVLSDGYAAHLGLAQVGSGRPAKG 1217  
RESULT 8  
09YV88 PRELIMINARY; PRT; 1250 AA.  
AC 09YV88  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE CG1111-PB.  
GN Name=rdgb; ORFNames=CG11111;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Mandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabot G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,  
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,  
RA Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,  
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Zhon M., Zhang G., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster".  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.



RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Githbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Paciel J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskaas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence,"  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kankiner J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.,  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomic perspective,"  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kankiner J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bencicourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.,  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review,"  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY051422; AAK92846.1; -;  
DR EMBL; AE003493; AAF48316.2; -;  
DR HSSP; P53810; IKCM.  
DR FLYBase; FBgn0003218; rdbp.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0016029; C:subrhabdomer cisterna; NAS.  
DR GO; GO:0008525; F:phosphatidylcholine transporter activity; IDA.  
DR GO; GO:0008526; F:phosphatidylinositol transporter activity; IDA.  
DR GO; GO:0016059; P:deactivation of rhodopsin mediated signaling; IMP.  
DR GO; GO:0007608; P:perception of smell; IMP.  
DR GO; GO:0007602; P:phototransduction; TAS.  
DR GO; GO:0016056; P:rhodopsin mediated signaling; IMP.  
DR InterPro; IPR004177; DDHD.  
DR InterPro; IPR001666; P\_transfer.  
DR Pfam; PF02862; DDHD; 1.  
DR Pfam; PF02121; IP\_trans; 1.  
DR PRINTS; PR00391; FTRANSFER.  
SQ SEQUENCE 1259 AA; 138894 MW; 9DD40B76EB1079F7 CRC64;  
Query Match 7.1%; Score 117.5; DB 2; Length 1259;  
Best Local Similarity 23.7%; Pred. No. 2.2;  
Matches 82; Conservative 39; Mismatches 106; Indels 119; Gaps 18;  
QY 14 FILSACGTLTGIPBHGGS-----KRAVEQELVNASARA---AYKD 51  
DB 906 FILRQIGKFGESIPVGSNDYKDNASFHQPREKIKRTSKLVKNAVNRADVTIVQE 965

QY 52 -----MDLQALHGRKVALYIATMGDGSGLT-----GGR--YSID- 85  
DB 966 GREORLNARFMYGLDMITLHGEKVDHI--MKPPRAGEWTFLESTVWDKNGRISYSID 1023  
QY 86 --ALIRGEYINSAPVRDYYT-----PRYETTAETTSGLTGLTSLIS-TLMAPAL 133  
DB 1024 QVSLGYGYIPVGMVVRGDHTSVDCYMAVPEPLTECVFISIDG---SFTASMSVGRDPKV 1080  
QY 134 SRTQSDGSGSSSLG---LNTGMDYRNET-----LTTNP--RD 168  
DB 1081 RAGAVDVCRRHMOELGYLLITYTGPDMQQQVWSWLSQHNFPHGLISFADGLSTDPGLGH 1140  
QY 169 TAPLSHVQVFFFRGIDVVSSPANADTVFINIDVFGTIRNRTEMLYNAAETLKAQTKE 228  
DB 1141 TAYLNNLVN---HGISITAYGSSND---ISVYTNVGRKTD-----Q 1177  
QY 229 YFVAVDRNKKLLIKPKTNAFEAAVKNYALMGPYKYSKIKPTEG 274  
DB 1178 IFIVGKVGKKL-----QSNNA--TVLSDGYAAHLAIGLVGVGSRPAKG 1217  
RESULT 10  
ID 08KG07 PRELIMINARY; PRT; 627 AA.  
AC 08KG07;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Alpha oxoglutarate ferredoxin oxidoreductase, alpha subunit.  
GN OrderedLocustNames=CT0163;  
OS Chlorobium tepidum.  
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
CX NCBI\_TaxID=1097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TLS / ATCC 49652 / DSM 12025;  
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;  
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,  
RA Dodson R.J., Deboy R.T., Gwin M.L., Nelson W.C., Haft D.H.,  
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,  
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,  
RA Parksey D.S., Niernan W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,  
RA Radune D., Vamathavan J.J., Khouiri H.M., White O., Gruber T.M.,  
RA Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.,  
RT "The complete genome sequence of *Chlorobium tepidum* TLS, a  
RT photosynthetic, anaerobic, green-sulfur bacterium,"  
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).  
DR EMBL; AE012796; AAM71411.1; -;  
DR TIGR; CT0163; -;  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR002869; POR.  
DR InterPro; IPR009014; Transketo\_C\_like.  
DR Pfam; PF01558; POR; 1.  
KW Complete proteome.  
SQ SEQUENCE 627 AA; 67379 MW; 4483729463DA2BA4 CRC64;  
Query Match 7.0%; Score 114.5; DB 2; Length 627;  
Best Local Similarity 21.6%; Pred. No. 1.5;  
Matches 61; Conservative 40; Mismatches 93; Indels 89; Gaps 12;  
QY 61 KVALYIATMGDGSGLTSGTGRYSIDLIRGEYNSAPVTRDYYPRYETTAETTSGLT 119  
DB 16 KTNVSVLPFADSGSGMQLTGTQFANVAVGSDLN-----TFPNPSSIRPAGVVA 67  
QY 120 GL-----TTSIST-----INAPLSR-----TQSDGSGSSSLGL 149  
DB 68 GVSQFQLOFGTGYTGAKEFDVMIAMNAAALKNLKHLHGIIIDTDGFDKKN--L 124  
QY 150 NIGMGDYRNETLTTNPRDPAFLSHLVQVTEFLAGIDVVSANADT-----DVFINID 202

DB 125 NLAYGE-----TNNPLEDGLTLD--YTVFKIPVISLRLQALADTGLSTKIIDRCKNMF 176  
QY 203 VFGTIRNRTMMLYNAETLKAEYFVAVDRTNKKLIKPTNAFEAY-----K 253  
DB 177 VLGV-----LYWLYSLPLETTIEALQSKFKNKQDIAENIKAVKAGYNGDETEMPS 228  
QY 254 ENYALMGPYKYSKGIKPTGELMVPESDIRPYGNHTGNSAPSV 296  
DB 229 QHGRECVPPAQKKGV-----YRRVTGNBASAI 256

## RESULT 11

08XSD6 PRELIMINARY: PRT: 3552 AA.  
ID 08XSD6  
AC 08XSD6  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)  
DE PROBABLE HEMAGGLUTININ-RELATED PROTEIN.  
CN Name=RS05701; OrderedLocustNames=RS0540;  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Plasmid megaplasmid.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GM11000;  
RX MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Broctier F., Camus J.C., Cactolico L.,  
RA Chandler M., Choiane N., Claudel-Bernard C., Cunac S., Demange N.,  
RA Gaopin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Sigler P., Thebault P., Whalen M., Winkler P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502(2002).  
DR EMBL: AL646079; CAD17691.1; -  
DR Pfam: PF05594; Fil\_haemagg; 21.  
DR Pfam: PF05860; Haemagg\_act; 1.  
DR SMART: SM00710; PBH1; 6.  
DR TIGRFAMs: TIGR01901; adhes\_NPGX; 1.  
DR TIGRFAMs: TIGR01731; fil\_hemag\_20aa; 15.  
DR PROSITE: PS01070; NUCLEASE\_NON\_SPEC; UNKNOWN\_1.  
KW Complete proteome; Plasmid.  
SQ SEQUENCE 3552 AA; 352934 MW; C5432AABE2CCF59C CRC64;

Query Match 6.8%; Score 111.5; DB 2; Length 3552;  
Best Local Similarity 20.0%; Pred. No. 26;  
Matches 84; Conservative 51; Mismatches 141; Indels 145; Gaps 17;

QY 16 LSACGLT-----GIPSHGGKRPFAVEQELVAASARAAYKMDLQALHGRVALYIATM 69  
DB 998 LSAANNLTLDGAGAVINQSGSMRGAVSLNNAIDNTSGKIGNDAGSGSVAMTTGSL 1057  
QY 70 GDGSG-----SLTGRYSIDA-----LIRGEYINSPARIDYTPRETTAE 112  
DB 1058 ANQGAISGDRNLSTTGLSDGRIIAGGCAITINGTYTHSANOIOAN-----NLTL 1113  
QY 113 TTSGGLTGLTSLSTLNAFALSRTOSDGS-----SRSSLGLNIG-----MGDYRETLTT 164  
DB 1114 TTTGLTNGT-LAAVNLTVAAVNDNAGADLMSASISVAGAGAITMAGIIBDITYTT 1172  
QY 165 -----NPRDTAFLSHLVQ-----TVFPLKGI 185  
DB 1173 QSGAFANTGVGNVNTLNAGASINTGASALAAATQNLVASDRLSNTGATLPSLGI 1232  
QY 186 DVVSPANADTVFINIDVGTIRNRTMMLYNAETLKAEYF--AVDRNKKLLIKP 243  
DB 1233 NIAANGARDN-----GLANRSNLVTNDOSTIEAGNLEIATQTLNTRPEPTVOT 1284  
QY 244 KTNAAEAYKE-----NYALMGPYKV-----SKGIX 270

DB 1285 VTTGTSTAHETRKGYIACATMNAAPHGCGTQAVWNNGYKXTPIDATFSTAQIVSQTSGRN 1344  
QY 271 PTEGLMV-----DESDIRPYGN-----HNGNAPSE---ADNSHSG 304  
DB 1345 PVDNVLVWVNGQNTIYNAVTNNGNGTVTVNYDADPH--NVVPSTEVATRSDDHNG 1403  
QY 305 Y 305  
DB 1404 Y 1404

## RESULT 12

06G428 PRELIMINARY: PRT: 874 AA.  
ID 06G428  
AC 06G428  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, last sequence update)  
DT 05-JUL-2004 (TREMBLrel. 27, last annotation update)  
DE Hypothetical protein.  
CN OrderedLocustNames=BH05510;  
OS Bartonella henselae (Rochalimaea henselae).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bartonellaceae; Bartonella.  
OX NCBI\_TaxID=38323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 49882 / Houston 1;  
RX PubMed=15210978; DOI=10.1073/pnas.0305659101;  
RA Alenaik U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardeli D.H.,  
RA Candaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvel M.,  
RA La Scola B., Holmberg M., Anderson S.G.E.;  
RT "The house-borne human pathogen Bartonella quintana is a genomic derivative of the zoonotic agent Bartonella henselae";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721(2004).  
DR EMBL: BX897699; CAF27359.1; -  
DR GO: GO:0007155; P:cell adhesion; IEA.  
DR InterPro: IPR006315; Autotransporter.  
DR InterPro: IPR005546; Auto\_transbeta.  
DR InterPro: IPR011050; Pectin\_lyas\_like.  
DR InterPro: IPR004899; Pectactin.  
DR InterPro: IPR003991; Pectactin\_C.  
DR Pfam: PF03797; Autotransporter; 1.  
DR Pfam: PF03212; Pectactin; 1.  
DR PRINTS: PR01484; PRTACTNFMALY.  
DR TIGRFAMs: TIGR01414; autotrans\_barl; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 874 AA; 95156 MW; 705381D8B17C958B CRC64;

Query Match 6.6%; Score 109; DB 2; Length 874;  
Best Local Similarity 19.9%; Pred. No. 6.5;  
Matches 65; Conservative 52; Mismatches 125; Indels 84; Gaps 13;

QY 11 FSVFLISAC-----GTLGIPSHGG-GKRFAVEQELVAASARAAYKMDL 54  
DB 9 FSVFLISCFVQVVASASENKRALKITSVTSKGKIAKSALVTGSLIPSSNIMIQDGV 68  
QY 55 QALHGRKALYIATMGDDGSGSLTGRYSIDLING-----EYINPAVATDVTYPRYET 110  
DB 69 EIVENGKTSIG-ATVEEGKQIVTRGKAIDTRKIGGKQFVEESPANFD--LORQSSA 125  
QY 111 AETSGGLTGLTSLSTL-NAPALSRTOSDGSRSSLGLNIGMGDYNETLTTNPRDT 169  
DB 126 YDAISGNGVGGQNVYDDAKAMNTKYTGSGEQLNLYMKNKKIKGGAFAVTVVSGNGRH 185  
QY 170 AFLSHL-VQVFFELGIDIVSP-----ANADTVFINIDVGTIRNRTMMLYNA 218  
DB 186 ILAEGMANNTILKDAVQVYVFGIIDLTVNGSASSMLYVGADLQGEIKVNDKGLY-- 243  
QY 219 ETLKQTKLEVFVAVDRN-----KLLIKPTNAFAEAYENVALMMPYKYSKGIKPTG 274  
DB 244 -----LFGRTTHIIRKLFVERGV-----AEWL-----FGV----- 271  
QY 275 LMVDFSDIRPYGNHTGNSAPVEADN 300

Db 272 -----GERNNETPOLEIDS 286

## RESULT 13

06BCF8 PRELIMINARY; PRT; 684 AA.

AC 06BCF8: 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Chitinase A.  
 GN Name=chitA;  
 OS Lyso bacter enzymes.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Lyso bacter.  
 NC NCB1\_TaxID=69;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3;  
 RA Kobayashi D.Y., Reedy R.M., Bick J., Yuen G.Y.;  
 RT "The chitA gene encodes for major chitinase activity expressed by  
 RT Lyso bacter enzymes C3."  
 RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 fibronectin type III domain.  
 DR EMBL: AY67480; AAT77163.1; -;  
 DR GO: GO:0005576; C:extracellular; IEA.  
 DR GO: GO:0030246; F:carbohydrate binding; IEA.  
 DR GO: GO:0016787; F:hydrolase activity; IEA.  
 DR GO: GO:0004553; F:hydrolase activity; IEA.  
 DR GO: GO:0005975; F:carbohydrate metabolism; IEA.  
 DR GO: GO:0008152; P:metabolism; IEA.  
 DR InterPro: IPR003610; CEM\_5\_12.  
 DR InterPro: IPR011583; Chitinase II.  
 DR InterPro: IPR003962; FNIII subd.  
 DR InterPro: IPR003961; FN\_III.  
 DR InterPro: IPR008957; FN\_III-like.  
 DR InterPro: IPR001223; Glyco\_hydro\_18.  
 DR InterPro: IPR001579; Glyco\_hydro\_18As.  
 DR Pfam: PF02839; CEM\_5\_12; 1.  
 DR Pfam: PF00041; FN3; 1.  
 DR Pfam: PF00704; Glyco\_hydro\_18; 1.  
 DR PRINTS: PR00014; FNTYPEIII.  
 DR ProDom: PD000471; Chitinase\_II; 1.  
 DR SMART: SM00495; ChEBD3; 1.  
 DR SMART: SM00650; FN3; 1.  
 DR SMART: SM00636; Glyco\_18; 1.  
 DR PROSITE: PS01095; CHITINASE\_18; 1.  
 DR PROSITE: PS50853; FN3; 1.  
 KW Glycosidase; Hydrolase.  
 SQ SEQUENCE 684 AA; 71453 MW; D54ECFE589E9E15E CRC64;

Query Match 6.5%; Score 107; DB 2; Length 684;

Best Local Similarity 26.1%; Pred. No. 6.8;

Matches 54; Conservative 25; Mismatches 84; Indels 44; Gaps 10;

QY 16 LSACGTLTG-----IPSHGGCKRAVGEOLVAARAARAVKMDQLHGK--VAL 64  
 DB 78 LGVCTGTGNGPPVSLTAPNSGATYAGAN---IAVSANAAADSDGSAVEFRGTTSL 134  
 QY 65 YIATMGDGS--GSLTGGYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGLT--- 119  
 DB 135 GVDTSAFPGVTWNNATAGNHSFTAVAKD---NQNAATTS-----SAVAATVVGSSDPT 185  
 QY 120 -----GTTSLSTLNAPALS-RTOSDSSGSRSSIGLNGMGDYRNTLTTPRDTAF- 171  
 DB 186 PPSVPGGLASSQTPANSVLSLWNASTDNSG-----GSGVAGYDYVRNGSLAGSPTTTSYT 240  
 QY 172 ---LSHLVQTFVFLRGIDVSPANADT 195  
 DB 241 VSGLSPTSTSYSTFVRARDNAGNAGS 267

## RESULT 14

06AVK7 PRELIMINARY; PRT; 731 AA.

AC 06AVK7: 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)  
 DE Putative transcription factor.  
 GN Name=OSUNBA0022C08.1;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 NC NCB1\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,  
 RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,  
 RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,  
 RA Vanaken S.S., Riedmuller S.B., Utecherback T.T., Feldblyum T.V.,  
 RA Wang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,  
 RA White O., Salzberg S.L., Fraser C.M.;  
 RT "Oryza sativa chromosome 3 BAC OSUNBA0022C08 genomic sequence."  
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Buell R.;  
 RL Submitted (AUG-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC097277; AAT81711.1; -;  
 DR InterPro: IPR005202; GRAS.  
 DR InterPro: IPR009014; Transketo\_C\_1like.  
 DR Pfam: PF03514; GRAS; 1.  
 SQ SEQUENCE 731 AA; 82218 MW; D3E04E65004364D7 CRC64;

Query Match 6.4%; Score 105.5; DB 2; Length 731;

Best Local Similarity 23.2%; Pred. No. 9.8;

Matches 63; Conservative 39; Mismatches 103; Indels 66; Gaps 12;

QY 93 INSPAVRTDYTYPRYETTAETTSGLTGLTSLTANAPALSRTOSSGSRSSIGLN-- 150  
 DB 134 LDSPESESTSYF-HSLASSTSNISGAVDS-----SQRRYGHSEYRLSGHSS 183  
 QY 151 ---IGMGDYRN-----ETLTTPRDTAFLSHLVQTFVFLRGIDVSPANADT-VFIN 200  
 DB 184 QPVPSPSDVYRNAMETLEDPLISNGRIPEYLFESFTWDFRGRVDEAQKFLPGSDKVID 243  
 QY 201 IDVFGITRNR-----TSMHLYNAETLKAQ-----TKLEFAYD----- 233  
 DB 244 LEAGVAKRODAGKAISLWKAELVKYKKNROSEDLDMEGRNSKOSAFCSDEPDWIE 303  
 QY 234 -----RTNKKLLIKPKYNAFEAAKYKENYALMGPYKV-SKGIKPTGLMVDVFSIDIRYG 286  
 DB 304 FDVLLRQTEKKATDLRGMRFEASKNGOVNQPKGSPGRSGRKPTK--KDVVDLRTLL 360  
 QY 287 NHTGNSAPSYEADNSHGGYGYSDAVRQHRQ 317  
 DB 361 IH---CAQAVAADDRRT---ANELLKQIRQ 384  
 RESULT 15  
 Q9SKF9 PRELIMINARY; PRT; 863 AA.  
 AC 09SKF9: 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Putative retroelement pol polyprotein.  
 GN Name=At2g12920;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NC NCB1\_TaxID=3702;



This Page Blank (uspto)



	Query Match	Similarity	Score	1645; DB 4;	Length	320;
	Best Local	Similarity	100.0%;	Pred.	2.3e-169;	
	Matches	320;	Conservative	0;	Mismatches	0;
					Indels	0;
					Gaps	0;
Qy	1	MRARLLIPLIFSVFLISACGTLTGIPSHGGGRFAVDEQLVAASARAAYKMDIDLQALHGR	60			
Db	1	MRARLLIPLIFSVFLISACGTLTGIPSHGGGRFAVDEQLVAASARAAYKMDIDLQALHGR	60			
Qy	61	KVALYATATGDDGSGSLTGGRYSIDALLRGEYINSPARTDYTPRYETTAETTSGLTG	120			
Db	61	KVALYATATGDDGSGSLTGGRYSIDALLRGEYINSPARTDYTPRYETTAETTSGLTG	120			
Qy	121	LITSLSLTNALPALSRTOQSDGSGSSLSGLNIGMGDIYNEFLTTNPPDITAFLSLVVTF	180			
Db	121	LITSLSLTNALPALSRTOQSDGSGSSLSGLNIGMGDIYNEFLTTNPPDITAFLSLVVTF	180			
Qy	181	FLRGDIVVSPAAADVDVFINIDVFSTINRTEEMHLYNAETLKAKDTKLEYFAVDRTNKKLL	240			
Db	181	FLRGDIVVSPAAADVDVFINIDVFSTINRTEEMHLYNAETLKAKDTKLEYFAVDRTNKKLL	240			
Qy	241	IKKRTNAFZAAIKENYALMGGPKYKSKIKETEGLMVDFSIIRRYGHTNGSNAPSVADN	300			
Db	241	IKKRTNAFZAAIKENYALMGGPKYKSKIKETEGLMVDFSIIRRYGHTNGSNAPSVADN	300			

Qy 301 SHEGYSDEAVRHRGQGP 320  
Db 301 SHEGYSDEAVRHRGQGP 320

RESULT 2  
US-09-043-302-7  
Sequence 7, Application US/09043302  
Patent No. 6617128  
GENERAL INFORMATION:  
APPLICANT: MEYER, Thomas F.  
APPLICANT: RUDEL, Thomas  
APPLICANT: SCHUEERFLOG, Ina  
APPLICANT: MAIER, Jurgen  
APPLICANT: EICKENJAGER, Sandra  
APPLICANT: SCHWAN, Thomas  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
TITLE OF INVENTION: the Adhesion of Neisseria Cells to Human Cells  
FILE REFERENCE: 0147-172P  
CURRENT APPLICATION NUMBER: US/09/043.302  
CURRENT FILING DATE: 1998-06-08  
EARLIER APPLICATION NUMBER: PCT/EP96/04092  
EARLIER FILING DATE: 1995-09-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 320  
TYPE: PRT  
ORGANISM: Neisseria gonorrhoeae  
US-09-043-302-7

Query Match 100.0%; Score 1645; DB 4; Length 320;  
Best Local Similarity 100.0%; Pred. No. 2.3e-169;  
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRARLLPIILFVFLIACGTLTGIPSHGGKRRFAVEDELYAASARA VKMDLQALHGR 60  
Db 1 MRARLLPIILFVFLIACGTLTGIPSHGGKRRFAVEDELYAASARA VKMDLQALHGR 60  
Qy 61 KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPARTTYTPRYVETTAETTSGLTG 120  
Db 61 KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPARTTYTPRYVETTAETTSGLTG 120  
Qy 121 LTTSLSTLNAPALSRTOGSGSRSSGLNIGMGDYNETLTTNPRDTAFLSHLVQTVF 180  
Db 121 LTTSLSTLNAPALSRTOGSGSRSSGLNIGMGDYNETLTTNPRDTAFLSHLVQTVF 180  
Qy 181 FLRGIDVSPNADTDVFINIDVGTINRRTTMMHLYNAETLKAQTKLEYFAVDRTNKKLL 240  
Db 181 FLRGIDVSPNADTDVFINIDVGTINRRTTMMHLYNAETLKAQTKLEYFAVDRTNKKLL 240  
Qy 241 IKPPTNAFEAAYKENYALMMGPYKSKGIKPTTEGLMVFSDIRPYNHTGNSAPSEADN 300  
Db 241 IKPPTNAFEAAYKENYALMMGPYKSKGIKPTTEGLMVFSDIRPYNHTGNSAPSEADN 300  
Qy 301 SHEGYSDEAVRHRGQGP 320  
Db 301 SHEGYSDEAVRHRGQGP 320

RESULT 3  
US-08-938-291A-9  
Sequence 9, Application US/08938291A  
Patent No. 6117673  
GENERAL INFORMATION:  
APPLICANT: Lev, Sima  
APPLICANT: PLOWMAN, Gregory D.  
APPLICANT: SCHLESINGER, Joseph  
TITLE OF INVENTION: RDGB PROTEINS AND RELATED  
TITLE OF INVENTION: PRODUCTS AND METHODS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: Lyon & Lyon

STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,291A  
FILING DATE: September 26, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/027,337  
FILING DATE: October 11, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Wardburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 228/172  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1250 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-938-291A-9

Query Match 7.1%; Score 117.5; DB 3; Length 1250;  
Best Local Similarity 23.7%; Pred. No. 0.0055;  
Matches 82; Conservative 39; Mismatches 106; Indels 119; Gaps 18;

Qy 14 FILSACGTLGIPSHGGG-----KRFVBOELVAASARA---AVKD 51  
Db 906 FILRQIGKFBGIPVGSNDKDNASFIHQGPREKRIKRSVKKVNAHNRADVQGE 965  
Qy 52 -----MDLQALHGRKVALYIATMGDQSGSLT-----GGR--YSID- 85  
Db 966 GREGRLNARFMYGFLDMITLHGEKVDVHL--MKDPAGQWTFLESTVTDKNGRISYSIPD 1023  
Qy 86 --AIRGEYINSPAVRDYTY-----PRVETTAETTSGLTGTLTSL--TLNAPAL 133  
Db 1024 QVSLGYSIYPVKVVRDHTSVDCYMAVVPPLTECVVFSIDG---SFTASVSVGRDPKV 1080  
Qy 134 SRTOGSGSRSSGL--LNIGMGDYNET-----LTTNP--RD 168  
Db 1081 RAGAVDVCRRHMOELGYLLIITGPRPDWQQRVSWLSQHPFPHGLISPADLSIDPLGHK 1140  
Qy 169 TAPLSHLVQTVFFLRGIDVSPNADTDVFINIDVGTINRRTTMMHLYNAETLKAQTKLE 228  
Db 1141 TAYLNVLVN---HGISITRAYGSSKD---ISVYTNVGKRTD-----Q 1177  
Qy 229 YFAVDRTNKKLLIKPTNAFEAAYKENYALMMGPYKSKGIKPTTEG 274  
Db 1178 IFIVGKVKKKL---QSNNA--TVLSDGYAAHLAQLQAVGSRPAKG 1217

RESULT 4  
US-09-589-619-9  
Sequence 9, Application US/09589619  
Patent No. 6576442  
GENERAL INFORMATION:  
APPLICANT: Lev, Sima  
APPLICANT: PLOWMAN, Gregory D.  
APPLICANT: SCHLESINGER, Joseph

```

RESULT 5
US-09-543-681A-4440
: Sequence 4440, Application US/09543681A
: Patent No. 6605709
: GENERAL INFORMATION:
: APPLICANT: GARY BRETON
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
: TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 2709.1002-001
: CURRENT APPLICATION NUMBER: US/09/543,681A
: PRIOR FILING DATE: 2000-04-05
: PRIOR APPLICATION NUMBER: US 60/128,706
: PRIOR FILING DATE: 1999-04-09
: NUMBER OF SEQ ID NOS: 8344
: SEQ ID NO 4440
: LENGTH: 190
: TYPE: PR1
: ORGANISM: Proteus mirabilis
US-09-543-681A-4440

```

```

Query Match          6.5%; Score 106.5; DB 4; Length 190;
Best Local Similarity 22.4%; Pred. No. 0.0039;
Matches 50; Conservative 34; Mismatches 84; Indels 39; Gaps 9;

QY      1 MRARLLPILEFSEVILSNACGTLTGIPSHGGKRFPAVEBELVAASRAAAVKMDLQALHGR 60
      6 LKKKMLVKALEFTVEVLISLGSVAASATINNGK- INFGEIVNAACVSTKSIDQIVNIGQ 64
QY      61 KVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGLTG 120
      65 YRTAQFDSVGC-----TVGNTDF-----YINLEACDT-----TVAQNASVSFSG 103
QY      121 LTTSLSTLMAPLSRQTSDGSGRSSGLINICMGMDYNEPLTTNPRDPAFLSHLVQTV- 179
      104 VSDS-NKRYTVAIVSNITTTGAGAAATGVGEI--TDHIGKVL--PPGDSVFSTAKQLID 156
QY      180 -----FFLR---GIDVYSPANAADTV 197
      157 GSNTLNFVARYKSLDVTVPGHADAV 183
Db

RESULT 6
US-09-489-039A-12770
Sequence 12770, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489.039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12770
LENGTH: 506
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12770

Query Match          6.5%; Score 106.5; DB 4; Length 506;
Best Local Similarity 22.8%; Pred. No. 0.019;
Matches 60; Conservative 39; Mismatches 85; Indels 79; Gaps 14;

QY      9 ILFSEVFLSACGT-----LTGIPS-HGGG-KRPAVEBELVAASRAAAVKMDLQALH-G 59
      257 VLTACYFTSSCVTGPAYARRLNGNESFPGTGMKECAFQOQ-----KKONDAPHIHG 307
QY      60 RKVALYIATMGDQSGSLTGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETT----- 114
      308 EKCWVYNAEM-----TTHVFEDYTTSPPIIRHD-GPRSNVYGVITVEKKP 351
Db

```

```

QY 115 -----SGGLTGLTSLSTLNAPALSRTOSSGSGS-----RSSLGINIGMGCD 156
DB 352 KNIIVSLVDGNPLYLIGSLISCSLSTNLSLCSRFQANSGSVFIRNPASLS-GGFD 410
QY 157 YRNSTLTTPRDTAFLS-HLVQTVFFLNGIDVSPANADTVFINIDVFGTIRNRTMHL 215
DB 411 EKEAVIHNLNAGESAFMDVTVLSNGVVSRL-----IFIAVDNNGMK-----I 453
QY 216 YNAETLKAQTKLE---YPAVRT 235
DB 454 TDTGMSPESEKLDKNDYVFDRS 476

```

```

RESULT 7
US-09-248-796A-15736
; Sequence 15736, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15736
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15736

```

```

Query Match 6.1%; Score 101; DB 4; Length 572;
Best Local Similarity 23.8%; Pred. No. 0.094;
Matches 65; Conservative 31; Mismatches 109; Indels 68; Gaps 12;
QY 64 LYIATMGDQSGSLTGRKYSIDALIRGYINSPAVRTTYTPRYETTRTSGGLTGLTT 123
DB 53 LSTTTNEDINLQSIDPTGVKIDC--SGAFIANDLLRD-TAP--EITDEINN-----S 100
QY 124 SLSTLNAPALSRTOSS-----DSSGSRSSSLGINIGMGDVARNLTTPRDTAFLSHLVQTV 179
DB 101 SSSSTGTPPLSNTSSDLEDDTEBSEFDEHNGGLSD---SSSTTVEDENQLOQLQOOP 157
QY 180 FF--LGGIDVVSANADTVFINIDVFGTIRNRTMHL YNAETLKAQTKLEYPADRTN 236
DB 158 LYGAVITKNNTITLPHQNDIHIISVDIGTL-----TKLVYFTKSSSN 200
QY 237 KKLIIKRTNA-----FEAAKYENVALMMGPYKVGSKIKPTBSLMDPFDIRP-- 284
DB 201 RKNSSRRTTGGGGGKGLHFPDFTENFENEMKMKMIKLQKSIK-----SNINPEI 252
QY 285 -YGNHTGNSAPSVADNSHEGYGSDAEVROHR 316
DB 253 TYIMATGGGA-----HKRYNLMKKTFKKHK 277

```

```

RESULT 8
US-09-489-039A-10635
; Sequence 10635, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29

```

```

; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10635
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10635
Query Match 5.9%; Score 97; DB 4; Length 515;
Best Local Similarity 20.5%; Pred. No. 0.21;
Matches 60; Conservative 42; Mismatches 130; Indels 60; Gaps 10;
QY 73 GSGSLTGRYSIDALIR-----GEYINSPA-----VRTDYTPRYETT--- 110
DB 145 GAGGYFTRPSLVMTIVYLIQPSLGEILQDPALISGSGFLVGSADSPFRNKYSHKEQYINPK 204
QY 111 -----ATTSGGLTGLTSLSTLNA-----PALSRTOSSGSGSRSSSLGINIGMG 155
DB 205 YQGVIEIKNTRRILCLMNTFLRLDANIIYGDALTDDAAELDKADVIVANPPFGNKGQOR 264
QY 156 DYRNSTLTTPN-RDTAFLSHLVQTVFFLNGIDVSPANADTVFINIDVFGTIRNRTMHL 214
DB 265 PLRSDIFPPTNNQALFLQHIYGLRAGGRAAVLPNNVLFESGVGTVDVARDLNNKCNLH 324
QY 215 YNAETLKAQTKLEYPADRTNKKLIIKPTNAPEAAKYENYA--LMMGPYKVGSKIKPT 272
DB 325 ----TILRLPTGLIFAHGVKTN--VLEFTKGTAKNKYQGNCTENWVYDLRTV----- 372
QY 273 EGLMVDSDIRPYGNHTGNSAPSVADNSHEG-----YGSDEAVRQROGO 319
DB 373 ---MPSFGKRTPGDADIGFAPDFTDPLHGAPEKYVGNPSTSKRTGE 421

```

```

RESULT 9
US-09-071-035-158
; Sequence 158, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```
US-09-071-035-158
Query Match      5.7%; Score 94.5; DB 4; Length 343;
Best Local Similarity 24.5%; Pred. No. 0.21;
Matches 53; Conservative 30; Mismatches 96; Indels 37; Gaps 10;

Qy 10 LFSVFTLSAC---GTLTGIPSHGCK-----RFAVQELVAAARAAMVMDLQALHG 59
Db 14 LFSVFTLSACGGGGSSSSSAGSGKAGEOVLRTGEOEMPTLSTATYKRISTIALNN 73

Qy 60 RKVALYIATMGD--QSGSGLTGGRYSIDALIRGEYIN-----SPAVRTDYTPRYET 109
Db 74 VYEGYIRLDKONKQVPAAGAEKAEVSEDLTYKILNDAKMSDCKPYTANDYVYGWQRT 133

Qy 110 TAETTSGLTGLTSLTSLTNAPALSRTOSSGSSSLGLNIGMGDYRNETLTTPRDT 169
Db 134 VDPATASEYAVLYASVK--NGDALIKGERD---KSELG--IKAVSD--TELEITLERAT 183

Qy 170 AFLSHLVQ-TVFPRLRGIDVVS-----PANADTVF 198
Db 184 PFYDYLAFPSFFPQRDIVERKYGNVANSBSAVY 219

RESULT 10
US-09-252-991A-17814
Sequence 17814, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17814
LENGTH: 496
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17814

Query Match      5.7%; Score 93.5; DB 4; Length 496;
Best Local Similarity 19.3%; Pred. No. 0.48;
Matches 89; Conservative 40; Mismatches 122; Indels 209; Gaps 21;

Qy 12 SVFTLSACGTLTGIPSHGCKRFAVEQELVAAARAAMVMDLQALHGRKVALYIATMGD 71
Db 52 SLFPRQAAGELS---EYGGARR--VEQDLTQALK-----QSLSKKKAKNVILLIGD 97

Qy 72 -QSGSGLT-----GGRY-SIDAL-IRGEYINSPAVRTDYTPRYETTAETTSGLT 119
Db 98 GMGSEIIVAAVYAGAGGYFKGIDALPLTQGYTHY-SLHKDSGLPDTVTSASAA---T 153

Qy 120 GLTTSLSLTNAP-----ALSRTOSSGSSRS----- 145
Db 154 AMTGVKSYNGAIGVDIHEQPHRNLBLAKLNGKATGNVSTAELODATPALLAHVTRAK 213

Qy 146 -----SLGLNIGMGDYRNETLTTPRDTALSHLVQVTFRLRGIDVVSPPANA 193
Db 214 CYGEPATSKQCPNSALENGAGSITEQWLKTRP-----DVLGGGA 254

Qy 194 DTDVFINIDVGTIRNRTMELYNAAETLKAQTKL-----EYFVDRTNKKLILKP 243
Db 255 -----ATFAETAKAGRYAKTILRAQAARAGYIVENLDELKAVRRANQK---QP 300

Qy 244 KTNAPFAAYKENVALMMP-----YKVSKGIK--- 270
Db 301 LIGLF--AFGNMPPVMIQPTATYHGNLQNPAPVSCBANPKRTADIPTLLAQMTSKAIELKLD 358
```

```
Qy 271 -----PTGLMVFSD--IRPYG-----NHTGNSAPSEVADNSH 302
Db 359 NPNGFPLQVEGASIDKQDHANPCGQIGETVYDLDEAVOKALAFKADSETLIVYADRAH 418

Qy 303 E-----GYGSDAVARQH 315
Db 419 SSQIIPETAPGLTQLLTTKDGAPLAISTYNSSESGGEH 458

RESULT 11
US-09-492-709A-302
Sequence 302, Application US/09492709A
Patent No. 6720139
GENERAL INFORMATION:
APPLICANT: Zyskind, Judith
APPLICANT: Ohlsen, Kari L.
APPLICANT: Trawick, John
APPLICANT: Foreyth, R. Allyn
APPLICANT: Froelich, Jamie M.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
FILE REFERENCE: ELITRA.001A
CURRENT APPLICATION NUMBER: US/09/492,709A
NUMBER OF SEQ ID NOS: 485
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 302
LENGTH: 2383
TYPE: PRT
ORGANISM: E. Coli
US-09-492-709A-302

Query Match      5.7%; Score 93; DB 4; Length 2383;
Best Local Similarity 23.5%; Pred. No. 7.2;
Matches 72; Conservative 43; Mismatches 102; Indels 90; Gaps 15;

Qy 31 GKRFAVQELVAAARAAMVMDLQALHGRKVALYIATMGDQSGSGLTGGRYSIDALIRG 90
Db 466 GKSGEVSLSVSLQTKYALKGYNEA-----TALEAA-----GKVVYTTGKDIIVLPAY 515

Qy 91 EYINSPAVRTDYTPRYETTAETTSGLTGLTSLTNAPALSRTOSSGSSSLGLN 150
Db 516 RFTSTP--ETDNTWP-LEVTADVYGNLSNREGSMVVQAFTLS--QKDSVSLSTQTLN 570

Qy 151 -----IGMGDYRNE-----TLT-----TNPRDTAFISH 174
Db 571 ADSSTATLTFTIADAGNPVYGLVLSRHEGVQDITLSDWKNGDSYTOILTTGAMSG 630

Qy 175 LVQTFVFLRGID-----VSPANADTVFINID-----VGTIRNRTMELHY 216
Db 631 TLTMPOLNGVDAKAPAVVVIISVSSRTHSSIKIKDRYLSGNPIEVTVELRDE---- 686

Qy 217 NAEITLKQTKLEFVAVRRTNKKLILKP-----KTNAPFAAYKENVALMMPRYKVSIGIK 270
Db 687 NDKPVKQKQQLNNAVSIDN---VKPGVTTDWKETA-DGVYKATYTAI-----TKSGG 735

Qy 271 PTGLMVF 277
Db 736 LTAKLIM 742

RESULT 12
US-09-841-786-3
Sequence 3, Application US/09841786
Patent No. 6669940
GENERAL INFORMATION:
APPLICANT: NAGARAJA, T. G.
APPLICANT: STEWART, GEORGE C.
APPLICANT: NARAYANAN, SANJEEV K.
APPLICANT: CHENGAPPA, M. M.
```

```

; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 927
; TYPE: PRT
; ORGANISM: Fusobacterium necrophorum
; US-09-841-786-3

```

Query Match 5.6%; Score 92.5; DB 4; Length 927;

Best Local Similarity 22.5%; Pred. No. 1.7;

Matches 80; Conservative 48; Mismatches 119; Indels 109; Gaps 19;

```

QY 11 FSVFLLACGTLTGIPSHGGKRAVEOE-----LVAASAR-AAVKMDLQALHGRKVAL 64
DB 558 FHTTSSGANGERDVSSVGAIVSM-VEQENYSKVSGKALAKDLNLIKAKETVN 616
QY 65 YIATMG---DQSGSLTGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTGL 121
DB 617 LVGNIGLARSTSGSAVGRLNVO---RSK--NSAIVEA-----KEKAE-----LSG- 658
QY 122 TTSISTNAPALST--QSDGSGRSSSLGNIIGMGDY----- 157
DB 659 ----ENINADALNRLFHVAGSPNGSGGNALNGSGYSGGISGARVSIDDEAYLKANKI 714
QY 158 ----RNETLTTPNPDYFL---SHLVQTFPLRGIDVSPA-----NADTD 196
DB 715 ALNSKDTSVNNAAGSAGICTKMAAVGAVAVNDYDISNKSISIDNDEGSKYDKNDDE 774
QY 197 VFINID-----VFGITRRTNTEHLYN-----AETLKAQTKLEYFAVDRNKKLLIK 242
DB 775 VTVVAESLEVDAKTGTGITSISVAGINKVSKPSEKPEERPEGFPGKIGNKVDSVK 834
QY 243 PK-TNAPEAAVKE--NYALMMGPYKSGIKPTGELMVDPSDIRPYGNHTGNSAPS 295
DB 835 NKITDSMDSLTEKITNY-----ISEGVKAGNLPNSVS-----HTPDKGPS 875

```

RESULT 13

US-09-841-786-1

; Sequence 1, Application US/09841786

; Patent No. 6669940

; GENERAL INFORMATION:

; APPLICANT: NAGARAJA, T. G.

; APPLICANT: STEWART, GEORGE C.

; APPLICANT: NARAYANAN, SANJEEV K.

; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN

; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF

; FILE REFERENCE: 30296

; CURRENT APPLICATION NUMBER: US/09/841,786

; CURRENT FILING DATE: 2001-04-24

; PRIOR APPLICATION NUMBER: 09/558,257

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 1

; LENGTH: 3241

; TYPE: PRT

; ORGANISM: Fusobacterium necrophorum

; US-09-841-786-1

Query Match 5.6%; Score 92.5; DB 4; Length 3241;

Best Local Similarity 22.5%; Pred. No. 13;

Matches 80; Conservative 48; Mismatches 119; Indels 109; Gaps 19;

QY 11 FSVFLLACGTLTGIPSHGGKRAVEOE-----LVAASAR-AAVKMDLQALHGRKVAL 64

```

DB 864 FHTTSSGANGERDVSSVGAIVSM-VEQENYSKVSGKALAKDLNLIKAKETVN 922
QY 65 YIATMG---DQSGSLTGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGLTGL 121
DB 923 LVGNIGLARSTSGSAVGRLNVO---RSK--NSAIVEA-----KEKAE-----LSG- 964
QY 122 TTSISTNAPALST--QSDGSGRSSSLGNIIGMGDY----- 157
DB 659 ----ENINADALNRLFHVAGSPNGSGGNALNGSGYSGGISGARVSIDDEAYLKANKI 1020
QY 158 ----RNETLTTPNPDYFL---SHLVQTFPLRGIDVSPA-----NADTD 196
DB 1021 ALNSKDTSVNNAAGSAGICTKMAAVGAVAVNDYDISNKSISIDNDEGSKYDKNDDE 1080
QY 197 VFINID-----VFGITRRTNTEHLYN-----AETLKAQTKLEYFAVDRNKKLLIK 242
DB 1081 VTVVAESLEVDAKTGTGITSISVAGINKVSKPSEKPEERPEGFPGKIGNKVDSVK 1140
QY 243 PK-TNAPEAAVKE--NYALMMGPYKSGIKPTGELMVDPSDIRPYGNHTGNSAPS 295
DB 1141 NKITDSMDSLTEKITNY-----ISEGVKAGNLPNSVS-----HTPDKGPS 1181

```

RESULT 14

US-09-489-039A-12983

; Sequence 12983, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 12983

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

; US-09-489-039A-12983

Query Match 5.6%; Score 91.5; DB 4; Length 766;

Best Local Similarity 24.4%; Pred. No. 1.6;

Matches 51; Conservative 33; Mismatches 72; Indels 53; Gaps 12;

```

QY 125 LSTINAPALSRQSDGSGRSSSLGLN-----IGMGDYRNETLTTPNPDYFLSHLVQTV 179
DB 374 IGTYSAPRLNKNKGDAIVGR--LDTNRITIDALISGMGGVRGDF-----NTGALSHKYNLG 425
QY 180 FFLR-GIDV---VSPNADTVFINIDV-----FG-----TIRNTEHLYNAE 219
DB 426 YAAQVHTDATTAKRSANPTTNIYDNDIVAMPDAAYGCVGHPLVSRSTQGLL-SD 484
QY 220 TLKAQTKLEYFAVDRNKKLLIKETNAF---EAYYENALMMGPYKSGIKPTGELM 276
DB 485 TLGFNDKVLFTAAARHQKVVVRVSNATGLEDTSSTYQSRWM-----FTGELV 534
QY 277 V-DESDIRPYGNHT-----GNSAPSVKAD 299
DB 535 YKPWEQLSVYANHTEALQPGSVADPTAAN 563

```

RESULT 15

US-09-463-402-2

; Sequence 2, Application US/09463402

; Patent No. 6596510

; GENERAL INFORMATION:

; APPLICANT: Lubitz, Werner

; APPLICANT: Resch, Stephanie

; TITLE OF INVENTION: Secretion of Carrier-bound Proteins into the Periplasm



This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 18, 2005, 22:52:34 ; Search time 111 Seconds  
(without alignments)  
1128.897 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645  
Sequence: 1 MRARLLIPILFVFLSACG.....SHEGYGDEAVRQRCOP 320

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10G_PUBCOMB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US10H_PUBCOMB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US11_PUBCOMB.pep.*
21:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
22:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1645	100.0	320	US-09-043-302-4	Sequence 4, Appl1
2	1645	100.0	320	US-09-043-302-7	Sequence 7, Appl1
3	1645	100.0	320	US-10-617-835-4	Sequence 4, Appl1
4	1645	100.0	320	US-10-617-835-7	Sequence 7, Appl1
5	1029	62.6	313	US-10-320-800-70	Sequence 70, Appl1
6	103.5	6.3	731	US-10-437-963-189840	Sequence 189840, A
7	101	6.1	329	US-10-425-114-38607	Sequence 38607, A
8	101	6.1	348	US-10-424-599-19578	Sequence 19578, A
9	101	6.1	445	US-10-282-122A-52671	Sequence 52671, A
10	100	6.1	3147	US-10-733-923-10253	Sequence 10253, A
11	100	6.1	3167	US-10-732-923-10252	Sequence 10252, A

12	99.5	6.0	245	15	US-10-424-599-254719	Sequence 254719, A
13	99	6.0	872	15	US-10-282-122A-55467	Sequence 55467, A
14	96.5	5.9	651	15	US-10-425-114-45324	Sequence 45324, A
15	95	5.8	685	15	US-10-335-977-4813	Sequence 4813, Ap
16	94.5	5.7	343	9	US-09-071-035-158	Sequence 158, App
17	94.5	5.7	343	14	US-10-206-576-158	Sequence 158, App
18	94.5	5.7	343	17	US-10-912-362-158	Sequence 158, App
19	94.5	5.7	525	15	US-10-425-114-49395	Sequence 49395, A
20	94.5	5.7	535	15	US-10-424-599-253649	Sequence 253649, A
21	94	5.7	431	15	US-10-282-122A-53853	Sequence 53853, A
22	93.5	5.7	764	17	US-10-732-923-23183	Sequence 23183, A
23	93.5	5.7	1074	15	US-10-282-122A-50616	Sequence 50616, A
24	93.5	5.7	1271	17	US-10-875-518-10	Sequence 10, Appl1
25	93.5	5.7	5836	15	US-10-378-083-20	Sequence 20, Appl1
26	93	5.7	451	15	US-10-282-122A-51597	Sequence 51597, A
27	93	5.7	764	17	US-10-732-923-23179	Sequence 23179, A
28	93	5.7	1000	14	US-10-156-761-13499	Sequence 13499, A
29	93	5.7	1325	9	US-09-741-669-304	Sequence 304, App
30	93	5.7	2383	9	US-09-912-020-302	Sequence 302, App
31	93	5.7	2383	15	US-10-282-122A-42852	Sequence 42852, A
32	93	5.7	2383	16	US-10-771-241-302	Sequence 302, App
33	92.5	5.6	893	16	US-10-637-544-4	Sequence 4, Appl1
34	92.5	5.6	893	17	US-10-819-275-4	Sequence 4, Appl1
35	92.5	5.6	927	9	US-09-841-786-3	Sequence 3, Appl1
36	92.5	5.6	927	15	US-10-647-057-3	Sequence 3, Appl1
37	92.5	5.6	3241	9	US-09-841-786-1	Sequence 1, Appl1
38	92.5	5.6	3241	15	US-10-647-057-1	Sequence 1, Appl1
39	92	5.6	245	16	US-10-425-115-250945	Sequence 250945, A
40	92	5.6	554	14	US-10-032-585-7607	Sequence 7607, Ap
41	92	5.6	1234	16	US-10-408-765A-155	Sequence 155, App
42	91.5	5.6	452	15	US-10-369-493-15597	Sequence 15597, A
43	91.5	5.6	452	15	US-10-369-493-15974	Sequence 15974, A
44	91.5	5.6	452	15	US-10-369-493-16336	Sequence 16336, A
45	91.5	5.6	523	15	US-10-282-122A-77204	Sequence 77204, A

# ALIGNMENTS

RESULT 1  
US-09-043-302-4 Application US/09043302  
Publication No. US20020086349A1  
GENERAL INFORMATION:  
APPLICANT: MEYER, Thomas F.  
APPLICANT: RUDEL, Thomas  
APPLICANT: SCHUEBERFLUG, Ina  
APPLICANT: MAIER, Jurgen  
APPLICANT: EICKERNJAGER, Sandra  
APPLICANT: SCHMAN, Thomas  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins which Impart  
TITLE OF INVENTION: The Adhesion of Neisseria Cells to Human Cells  
FILE REFERENCE: 0147-172P  
CURRENT APPLICATION NUMBER: US/09/043,302  
CURRENT FILING DATE: 1998-06-08  
EARLIER APPLICATION NUMBER: PCT/EP96/04092  
EARLIER FILING DATE: 1995-09-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 320  
TYPE: PRT  
ORGANISM: Neisseria gonorrhoeae  
US-09-043-302-4

Query Match 100.0%; Score 1645; DB 9; Length 320;  
Best Local Similarity 100.0%; Pred. No. 1.8e-150;  
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1 MRARLLIPILFVFLSACGTTGTPSHGGGRFVVEOBELVAASRAAVKMDLQALHGR 60  
Db 1 MRARLLIPILFVFLSACGTTGTPSHGGGRFVVEOBELVAASRAAVKMDLQALHGR 60

```

Qy      61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
        61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
Db      121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
        121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
Qy      121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
        121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
Db      181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
        181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
Qy      181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
        181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
Db      241 IKPRTNAFEAAYKENYALMGPYKVS KGIKPTGELMVD FSDIRPYGNHTGNSAPSEADN 300
        241 IKPRTNAFEAAYKENYALMGPYKVS KGIKPTGELMVD FSDIRPYGNHTGNSAPSEADN 300
Qy      301 SHEGYSDEAVRQHRQGP 320
        301 SHEGYSDEAVRQHRQGP 320
Db      301 SHEGYSDEAVRQHRQGP 320

```

## RESULT 2

```

US-09-043-302-7
; Sequence 7, Application US/09043302
; Publication No. US20020086349A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHEUERPFUG, Ina
; APPLICANT: MAIER, Jurgen
; APPLICANT: EICKERNJAGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; FILE REFERENCE: 0147-172P
; CURRENT APPLICATION NUMBER: US/09/043,302
; EARLIER FILING DATE: 1998-06-08
; EARLIER APPLICATION NUMBER: PCT/EP96/04092
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-043-302-7

```

```

Query Match      100.0%; Score 1645; DB 9; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MRARLLPIILFSVFLISACGTLTGIPSHGGGRFAVEQELVAASARAAYKMDIQALHGR 60
        1 MRARLLPIILFSVFLISACGTLTGIPSHGGGRFAVEQELVAASARAAYKMDIQALHGR 60
Db      61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
        61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
Qy      61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
        61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
Db      121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
        121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
Qy      121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
        121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
Db      181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
        181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
Qy      181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
        181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
Db      241 IKPRTNAFEAAYKENYALMGPYKVS KGIKPTGELMVD FSDIRPYGNHTGNSAPSEADN 300
        241 IKPRTNAFEAAYKENYALMGPYKVS KGIKPTGELMVD FSDIRPYGNHTGNSAPSEADN 300
Qy      301 SHEGYSDEAVRQHRQGP 320
        301 SHEGYSDEAVRQHRQGP 320

```

```

Db      301 SHEGYSDEAVRQHRQGP 320

```

## RESULT 3

```

US-10-617-835-4
; Sequence 4, Application US/10617835
; Publication No. US20050124037A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHEUERPFUG, Ina
; APPLICANT: MAIER, Jurgen
; APPLICANT: EICKERNJAGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; FILE REFERENCE: 0147-172P
; CURRENT APPLICATION NUMBER: US/10/617,835
; EARLIER FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US/09/043,302
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: PCT/EP96/04092
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-617-835-4

```

```

Query Match      100.0%; Score 1645; DB 18; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MRARLLPIILFSVFLISACGTLTGIPSHGGGRFAVEQELVAASARAAYKMDIQALHGR 60
        1 MRARLLPIILFSVFLISACGTLTGIPSHGGGRFAVEQELVAASARAAYKMDIQALHGR 60
Db      61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
        61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
Qy      61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
        61 KVALYIATMGDGGSLTGGRYSIDALIRGEYINSPAVRTDYTYRYETTAETTSGLTIG 120
Db      121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
        121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
Qy      121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
        121 LTTSLSTLNAPALSRITQSDGSSRSLGINTGMDYRNETLTTPRDTAFLSHVQTVF 180
Db      181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
        181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
Qy      181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
        181 FLRGIDVVS PANADTDVFINIDVFGTIRNRTTEMLYNAETLKAQTKLEYFAVDRTNKKLL 240
Db      241 IKPRTNAFEAAYKENYALMGPYKVS KGIKPTGELMVD FSDIRPYGNHTGNSAPSEADN 300
        241 IKPRTNAFEAAYKENYALMGPYKVS KGIKPTGELMVD FSDIRPYGNHTGNSAPSEADN 300
Qy      301 SHEGYSDEAVRQHRQGP 320
        301 SHEGYSDEAVRQHRQGP 320

```

## RESULT 4

```

US-10-617-835-7
; Sequence 7, Application US/10617835
; Publication No. US20050124037A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHEUERPFUG, Ina
; APPLICANT: MAIER, Jurgen
; APPLICANT: EICKERNJAGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; FILE REFERENCE: 0147-172P
; CURRENT APPLICATION NUMBER: US/10/617,835
; EARLIER FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US/09/043,302
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: PCT/EP96/04092
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-617-835-7

```

FILE REFERENCE: 0147-172P  
 CURRENT APPLICATION NUMBER: US/10/617,835  
 CURRENT FILING DATE: 2003-07-14  
 PRIOR APPLICATION NUMBER: US/09/043,302  
 PRIOR FILING DATE: 1998-06-08  
 PRIOR APPLICATION NUMBER: PCT/EP96/04092  
 PRIOR FILING DATE: 1995-09-18  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 7  
 LENGTH: 320  
 TYPE: PRT  
 ORGANISM: *Neisseria gonorrhoeae*  
 US-10-617-835-7

Query Match 100.0%; Score 1645; DB 18; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 1,88-150;  
 Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRARLLILPILSVFLISACGTLTGIPSHGGKRAVEQELVAASRAAVKMDLQALHGR 60  
 DB 1 MRARLLILPILSVFLISACGTLTGIPSHGGKRAVEQELVAASRAAVKMDLQALHGR 60  
 QY 61 KVALYATMGDGGSLTGGGRYSIDALIRGEYINSPAVRTDVTYPRVETTAETTSGLTGG 120  
 DB 61 KVALYATMGDGGSLTGGGRYSIDALIRGEYINSPAVRTDVTYPRVETTAETTSGLTGG 120  
 QY 121 LTTSLTNAPALSRTOGSGSRSSLGINIGMDYRENTLTTPRDTAFSLHVOYVF 180  
 DB 121 LTTSLTNAPALSRTOGSGSRSSLGINIGMDYRENTLTTPRDTAFSLHVOYVF 180  
 QY 181 FLRGIDVSPANADTVFINIDVFETIRNRTMHLNNAETLKAQTKLEYFAVDRTNKKLL 240  
 DB 181 FLRGIDVSPANADTVFINIDVFETIRNRTMHLNNAETLKAQTKLEYFAVDRTNKKLL 240  
 QY 241 IKPKTNAFEAAKYENYALMGPYKYSKGIKPTGELVWPSPDRPFGNHTGNSAPSEADN 300  
 DB 241 IKPKTNAFEAAKYENYALMGPYKYSKGIKPTGELVWPSPDRPFGNHTGNSAPSEADN 300  
 QY 301 SHEGYGSDAENVQRHQQGP 320  
 DB 301 SHEGYGSDAENVQRHQQGP 320

RESULT 5  
 US-10-320-800-70  
 Sequence 70, Application US/10320800  
 Publication No. US20030215469A1  
 GENERAL INFORMATION:  
 APPLICANT: ROBINSON, ANDREW  
 APPLICANT: GORRINGE, ANDREW  
 APPLICANT: HUDSON, MICHAEL  
 APPLICANT: REDDIN, KAREN  
 TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE  
 FILE REFERENCE: 1581.0790001  
 CURRENT APPLICATION NUMBER: US/10/320,800  
 CURRENT FILING DATE: 2002-12-17  
 PRIOR APPLICATION NUMBER: PCT/GB99/03626  
 PRIOR FILING DATE: 1999-11-02  
 NUMBER OF SEQ ID NOS: 75  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 70  
 LENGTH: 313  
 TYPE: PRT  
 ORGANISM: *Neisseria meningitidis*  
 US-10-320-800-70

Query Match 62.6%; Score 1029; DB 15; Length 313;  
 Best Local Similarity 62.7%; Pred. No. 8,3e-91;  
 Matches 197; Conservative 49; Mismatches 62; Indels 6; Gaps 2;  
 QY 5 LILPILSVFLISACGTLTGIPSHGGKRAVEQELVAASRAAVKMDLQALHGRKVAL 64  
 DB 5 LILPILSVFLISACGTLTGIPSHGGKRAVEQELVAASRAAVKMDLQALHGRKVAL 64

DB 6 LILPL-----VLTACGTLTGIPAHGGKRAVEQELVAASRAAVKEMDLSLKGKRAAL 60  
 QY 65 YIATMGDGGSLTGGGRYSIDALIRGEYINSPAVRTDVTYPRVETTAETTSGLTGLTSS 124  
 DB 61 YSVWVGDDGSSNIGSGRYSIDALIRGHNHPESATQSYPAVDYATTKSDALSSVTS 120  
 QY 125 LSTLNAPALSRTOGSGSRSSLGINIGMDYRENTLTTPRDTAFSLHVOYVFPRG 184  
 DB 121 TSLNAPALSRTOGSGSRSSLGINIGMDYRENTLTTPRDTAFSLHVOYVFPRG 180  
 QY 185 IDVSPANADTVFINIDVFETIRNRTMHLNNAETLKAQTKLEYFAVDRTNKKLLIKPK 244  
 DB 181 IEVPEPEYADTVFINIDVFETIRNRTMHLNNAETLKAQTKLEYFAVDRTNKKLLIKPK 240  
 QY 245 TNAPFAAKENYALMGPYKYSKGIKPTGELVWPSPDRPFGNHTGNSAPSEADNHEG 304  
 DB 241 TAAVESQYQDEQYALMTGTYKYSKTVKASDRMWPFSDITPYGDTTAQRPFKQNGKPK 300  
 QY 305 YGSDAENVQRHQQGP 318  
 DB 301 -DVGNEVIRRRKGG 313

RESULT 6  
 US-10-437-963-189840  
 Sequence 189840, Application US/10437963  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Barbazuk, Brad  
 TITLE OF INVENTION: RICE Nucleic Acid Molecules and Other Molecules Associated With  
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(53221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 204966  
 SEQ ID NO 189840  
 LENGTH: 731  
 TYPE: PRT  
 ORGANISM: *Oryza sativa*  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_8630C.1.pep  
 US-10-437-963-189840

Query Match 6.3%; Score 103.5; DB 16; Length 731;  
 Best Local Similarity 23.2%; Pred. No. 1.3;  
 Matches 63; Conservative 39; Mismatches 103; Indels 66; Gaps 12;  
 QY 93 INSPAVRTDVTYPRVETTAETTSGLTGLTSLTNAPALSRTOGSGSRSSLGIN-- 150  
 DB 134 LDBSESTSSYP-HSLASVTSNISCANVS-----SQRYVGHSEYRSLSGSS 183  
 QY 151 ---IGMDYRN-----ETLTTPRDTAFSLHVOYVFPRGIDVSPANADTVFIN 200  
 DB 184 QPVPSSDVSNNAMETLBDPLISNGRIPEYLFESPTDFRQVDAQKFLPGSDKVID 243  
 QY 201 IDVFETIRN-----TEHNLNNAETLKAQ-----TKLEYFAVD----- 233  
 DB 244 LEAGVAKROBAGKASISLNVSKAEVLKVKURQSEDLDMEGRNSKQSAFCSDEPDWTEM 303  
 QY 234 -----RTNKKLLIKPKTNAFEAAKYENYALMGPYKYSKGIKPTGELVWPSPDRPFG 286  
 DB 304 FDLVLRQTEKATDURKMMRFASNSPVAPKPGSGSRSGRKRKT--KDVVDLRTLL 360  
 QY 287 NHTGNSAPSEADNHSHEGYGSDAENVQRHQQGP 317  
 DB 361 IH---CAQVAADDRRT---ANELKKQIRQ 384

```

Db      91 GTLANVVGCGCRNKRVKRP.LITTNSSA-----AIDTAASNNSSNSSA 136
QY      79 GGRYSIDALLRGEYINS-----PAVRTDYVPYRYETAAE--TTSGGLTGTLTSLST LMA 130
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      137 PLQPIIDTASTNSNININPLFYGLPSSSSDVNLPLFRRPGRSISSGFDLQJLNNALGLGFSS 196
QY      131 PALSRTOSSDGSGRSSSLGLNIGMGDYRNETLTTNPEDTAFI..SHLVQTVFPLRGID 186
           : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      197 GVLSNEASDNNNGRYNMGFSSNNTLISSTTSTTTTPAMSRLLSSSLQCKFMTDGVD 253

```

/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 / FILER INFORMATION: 60/206,848  
 / FILE REFERENCE: ELITRA.034A  
 / CURRENT APPLICATION NUMBER: US/10/282,122A  
 / CURRENT FILING DATE: 2003-02-20  
 / PRIOR APPLICATION NUMBER: 60/191,078  
 / PRIOR FILING DATE: 2000-03-21  
 / PRIOR APPLICATION NUMBER: 60/206,848  
 / PRIOR FILING DATE: 2000-05-23  
 / PRIOR APPLICATION NUMBER: 60/207,727  
 / PRIOR FILING DATE: 2000-05-26  
 / PRIOR APPLICATION NUMBER: 60/230,335  
 / PRIOR FILING DATE: 2000-09-06  
 / PRIOR APPLICATION NUMBER: 60/230,347  
 / PRIOR FILING DATE: 2000-09-09

```

PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/267,636
PRIORITY FILING DATE: 2001-02-09
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52671
LENGTH: 445
TYPE: PRT
ORGANISM: Clostridium botulinum
US-10-282-122A-52671

Query Match
Beet Local Similarity 6.1%; Score 101; DB 15; Length 445;
Matches 73; Conservative 40; Mismatches 82; Gaps 15;

62 VALYIATMGDSG-----SLTGRYSIDALRGEYINSPAVRTDLYPRRYETTAETTS 115

```

Query Match	6.1%	Score 101,	DB 15;	Length 445;
Best Local Similarity	23.5%;	Pred.	No. 1.1;	
Matches	73;	Conservative	40;	Mismatches 82; Gaps 15;
Oy	62	VALLIATMDGSGS-----SLTGRYSIDALIRGEYINSPAVRTTYPRYETTAETTS	115	
Db	45	IAPVFLI---GGGVFLVLYSVTGKMAIIALVNKGILNLSLDSKLEMLIEKRLLVVG	101	
Oy	116	-----CGTLGLTSL-----FLMAPLRSTQSDSSG---RSISGLNIGMGXYRNRT	161	
Db	102	PKIVAGGGGTGLSTMRLGKRYRYSNTAITVAADDGGGGSGELKEBDLGML--PPGDIRNCI	159	

```

NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 10252
; LENGTH: 3167
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-732-923-10252

Query Match 6.1%; Score 100; DB 17; Length 3167;
Best Local Similarity 22.7%; Pred. No. 24;
Matches 58; Conservative 40; Mismatches 94; Indels 64; Gaps 11;

OY TMGDSGSGSLTGG----RYG-----IDLIRGEVINSIPA-----VRTDYTPRYE 108
Db TIGSEBDSGWSGSDRGRRSRITTDQVYIDAFQIGIPDSQSTSTAHKPPRTIETHDDY 2536
OY TTA-----BTSGGLTGLTSLSTL-----NAPALSRQS-----DGSGRSSL 147
Db TITSLQEDLDATDSEVDGEMLDVSTFPDDILGKSMDEAFLSTKSLREHTDTSIDRKS 2596
OY 148 GLNIGKMGDYRNELT-----TNPRDTAFSLVQVYFPLRGIDVSPNADTD 196
Db 2597 GEKHSY--YRNRTDTSIDRKSREVITTEDELTSELODEIMKLVF-----VEPSVSKSD 2648
OY 197 VFINIDVFETIRNTEWHLYNAETLKQOTLEYFAVDRTKKLLIK--PKTNAFEAAVKE 254
Db 2649 SSANIK--ASQNSTTKPCDEBELLEIKSEYFLIKGYSLLIPKSDPLGKMLQKLRQ 2705
OY 255 NYALMMGPYKVSQIK 270
Db 2706 NDLSLATLDFSLTRKX 2721

RESULT 12
US-10-424-599-254719
; Sequence 254719, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254719
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72033C.1.pep
US-10-424-599-254719

Query Match 6.0%; Score 99.5; DB 15; Length 245;
Best Local Similarity 27.4%; Pred. No. 0.63;
Matches 51; Conservative 19; Mismatches 69; Indels 47; Gaps 8;

OY 23 TGIPSHGGGRPAVEGEVLVAASRAAVKMDLQALHGRKALYIATM-----GQGGSG 75
Db 77 TGRSRGFVTFATSEB-----ASSALQMGDQDLGRRIRRVAVATERSRPGFGDGGYG 131
OY 76 SLNCGRSIDALIRGEVINSPAVRTDYTPRYETTAETTSGLTGLTSLSTANAPALSR 135
Db 132 SGSGSGGNNRG-----GNYSGGGYSSGSGYK--GNYSGCGY-NVTSYSSGNA-----E 179
OY 136 TQSDSGSGRSS--LGLNIGMGDYRNELTNTNPRDTAFSLVQVYFPLRGIDVSPANA 193
Db 180 TSYTGGAASNYQFENSGGVFGSASGEFSSNDT-----AGA 218
OY 194 DTDVFI 199

```

[illegible]

```

Qy      276 -----MVFSDIRPYGNTGNSAP-----SVEADNSHEGCGYSD 309
          : | | | | | | | | | | | | | | | | | | | | | |
Db      348 LIGSSATLMAFYWLMTYGSGDLSPRSAFELLAADANAAIDVRSEEMREKXGIPD 404

```

5  
5-977-4813  
Office 4813, Application US/10335977  
Patent No. US2004005279A1  
PATENT INFORMATION:  
APPLICANT: DOUGLAS SMITH et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: Windows NT 4.0

SOFTWARE: UNIX

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/335,977

FILING DATE: 30-Dec-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/993,002

FILING DATE: 17-Dec-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandagouras, Amy E.

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: GTN-018

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 4813:

SEQUENCE CHARACTERISTICS:

LENGTH: 685 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Helicobacter pylori

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (B) LOCATION 1...685

SEQUENCE DESCRIPTION: SEQ ID NO: 4813:

US-10-335-977-4813

Query Match 5.8%; Score 95; DB 15; Length 685;

Best Local Similarity 20.8%; Pred. No. 7.8;

Matches 64; Conservative 43; Mismatches 126; Indels 74; Gaps 13;

```
QY 30 GGRFAVQELVAASARAANKMDIQ---ALHGRVALYIATMGDQSGSLTGRYSIDA 86
DB 309 GFSQTTQESDLKKSAR-TIKGKLEGVVSLNGOKLDLSALTYESNTSENT-----DA 361
QY 87 LIRGEYINSPAVRTDYTPRYETTAETTSGLTGLTTSISTLUNAPALSRTOGSGSRSS 146
DB 362 IIGA--INAKGELNAFKNAEGKLVINSKGMJT-----IKGEDALGKASLKD 406
QY 147 LGLNIGMGMDYRNETLTNPRDTAFSLHVQ---TVFFLRGIDVVSANADTVFINID 202
DB 407 LGLSAGMVOSYEAS-----ODTLFMSKVLQKASDSQFTYNGVSIITRPNEVNDVISGVN 460
QY 203 VEGTIRNRTEHMLVNAETLKAKQTKLEYFAVDRTNKKLIIKPTNAFEAAVKE----- 254
DB 461 I--TIEQTE-----PNKPAIISVSHDNQAIIDSLE--EFVKAHNEELPKLDED 505
QY 255 -NYALMGPYKYSKGIKPTGELMVDPSDIPRYGNHTGNSAPSEVADNSHEGYGS----- 308
DB 506 TRYDADTKIAGIFNGVDIRAIRSSLNNVFSYVHTDNGVESLMK-----YGLSLDDKG 559
QY 309 ----DEA 311
DB 560 VMSLDEA 566
```

Search completed: August 18, 2005, 22:56:48  
Job time : 116 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 18, 2005, 22:54:53 ; Search time 3940 Seconds  
(without alignment)

3935.449 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645

Sequence: 1 MRALLPILFSLVFLSNAG.....SHEGYGYSDEAVRHQRCOP 320

## Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=frame\_plus.p2n.model -DEV=xlp  
-Q=/cgn2.1/USPRO.spool.p/US10617835/runcat.18082005.115607.8361/app.query.fasta.1.519  
-DB=genembl -OPMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biolum62 -TRANS=human0.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTWT=PCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10617835 @CGN 1.1.5600 @runcat.18082005.115607.8361 -NCPU=6 -ICPU=3  
-NO.MMAP -LARGESQUERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_ov.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_pro.\*  
11: gb\_scs.\*  
12: gb\_ey.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1645	100.0	1136	6	A61825 Sequence 5
2	1645	100.0	1136	6	AR393746 Sequence 1
3	1645	100.0	3287	6	A61821 Sequence 1
4	1645	100.0	3287	6	AR393745 Sequence

5	1645	100.0	3294	1	AF142582	AF142582 Neisseria
6	1641	99.8	963	6	A96050	A96050 Sequence 83
7	1641	99.8	963	6	AX04397	AX04397 Sequence
8	1634	99.3	963	6	A96046	A96046 Sequence 79
9	1634	99.3	963	6	AX043995	AX043995 Sequence
10	1624	98.7	963	6	A96048	A96048 Sequence 81
11	1624	98.7	963	6	AX043999	AX043999 Sequence
12	1624	98.7	340806	1	NMA122491	AL162752 Neisseria
13	1621	98.5	10869	1	AE002559	AE002559 Neisseria
14	1621	98.5	172325	6	AX044035	AX044035 Sequence
15	1621	98.5	942	6	A96284	A96284 Sequence 31
16	1620	98.5	640	6	A96044	A96044 Sequence 77
17	1620	98.5	942	6	A96280	A96280 Sequence 31
18	1620	98.5	10057	1	AE002420	AE002420 Neisseria
19	1620	98.5	11700	1	AE002393	AE002393 Neisseria
20	1620	98.5	349980	6	AX044029	AX044029 Sequence
21	1620	98.5	349980	6	AX044030	AX044030 Sequence
22	1626	62.4	195767	1	NMA722491	AL162758 Neisseria
23	1022	62.1	942	6	A96282	A96282 Sequence 31
24	650	39.5	591	6	A96278	A96278 Sequence 31
25	122.5	7.4	2233	1	AY069934	AY069934 Mycobacte
26	119.5	7.3	2723	1	A1667480	A1667480 Lyobacte
27	117.5	7.1	3418	3	BT011130	BT011130 Drosophi
28	117.5	7.1	4188	3	AY051422	AY051422 Drosophi
29	117.5	7.1	4276	6	CO595557	CO595557 Sequence
30	117.5	7.1	4300	6	CO595443	CO595443 Sequence
31	117.5	7.1	4310	3	DMR684PT	Y08035 D.melanogas
32	114.5	6.8	12993	1	AE012796	AE012796 Chlorobiu
33	114.5	6.8	5529	6	BD262969	BD262969 DNA seq
34	111.5	6.8	5529	6	AX024260	AX024260 Sequence
35	111.5	6.8	47713	6	BD262937	BD262937 DNA seq
36	111.5	6.8	47713	6	AX024213	AX024213 Sequence
37	111.5	6.8	213050	1	AL646079	AL646079 Ralstonia
38	109	6.6	110000	1	EX897699	Continuation (7 of
39	107.5	6.5	4078	1	AB007125	AB007125 Seretia
40	107.5	6.5	299925	1	AP005048	AP005048 Streptomy
41	106.5	6.5	573	6	AR375262	AR375262 Sequence
42	106.5	6.5	573	6	AR388870	AR388870 Sequence
43	105.5	6.4	2196	6	AX652816	AX652816 Sequence
44	105.5	6.4	2902	8	AK101402	AK101402 Oryza sat
45	105.5	6.4	3421	8	AK101451	AK101451 Oryza sat

## ALIGNMENTS

RESULT 1	A61825	Sequence 5 from Patent WO9711181.	1136 bp	DNA	linear	PAT 09-MAR-1998
LOCUS	A61825					
DEFINITION	A61825					
ACCESSION	A61825					
VERSION	A61825.1	GI:3715996				
KEYWORDS						
SOURCE						
ORGANISM	Neisseria gonorrhoeae					
	Neisseria gonorrhoeae					
	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;					
	Neisseriaceae; Neisseria.					
REFERENCE						
AUTHORS	Meyer, T.F., Rudel, T., Scheuerpflug, I., Fischer, Eckhard and					
	Maler, J.					
TITLE	NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE					
	ADHESION OF NEISSERIA CELLS TO HUMAN CELLS					
JOURNAL	Patent: WO 9711181-A 5 27-MAR-1997;					
	MAX PLANCK GEBELTSCHAFT (DE)					
COMMENT	Other Publication DE 19534579 970320.					
FEATURES						
source	1..1136					
	/organism="Neisseria gonorrhoeae"					
	/mol_type="unassigned DNA"					
	/strain="MS11"					
	/db_xref="taxon:485"					
ORIGIN						
Alignment Scores:						

Pred. No.: 3.68e-124 Length: 1136  
 Score: 1645.00 Matches: 320  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x A61825 (1-1136)

QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 DB 135 ATGGGGGACGGCGCTGATACCTAATTCCTTTTGGTTTAAATTATCCGCTGGCGG 194  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyIleAspPheAlaValGlnGlnLeu 40  
 DB 195 ACACGACAGGATTCATCCGATGCGGAGGCAAAACGCTTCGGCTGGAACAGAACTT 254  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValIleAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 DB 255 GTGGCCGCTTCGCGAGAGCTGCCGTTAAAGCATGATTTACAGCATTTACACGACGA 314  
 QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGly 80  
 DB 315 AAAGTTGATTGATTCATTGCAACTATGAGCGACCAAGGTTCAAGGATTTGACAGGGGT 374  
 QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
 DB 375 CGCTACTCATTTGATGACCTGATTCGCGCGAATCATTAACAGCCCTGCGCTCGGCACC 434  
 QY 101 AspTyrThrTyrProArgTyrGlyLeuThrAlaGlnThrSerGlyLeuThrGly 120  
 DB 435 GATTACACCTTATCCGCTTACGAAACCAACCGCTGAAACAAATCATGAGCGTTGACGGGT 494  
 QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 DB 495 TTAACTCATCTTTATCTAATCACTTAATGCCCTGACCTCGCGACCAATTCAGACGT 554  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
 DB 555 AGCGGAATAGAGAGAGCTGGGCTTAATATTTGGCGGAGTGGGGATTTATCGAAATGAA 614  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 DB 615 ACCTTGACGACCAACCCCGCGACACTGCTTTCTTCCACTTGGTACAGACCGTATTT 674  
 QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 DB 675 TTCTGCGCGGATAGACGTTGTTTCTCTGCCAATGCCGATACAGATGTTTATTAAAC 734  
 QY 201 IleAspValPheGlyThrIleAsnAsnArgThrGluMetHisLeuTyrAsnAlaGlnThr 220  
 DB 735 ATCGACGATTCGGAGACGATACGCAACGAAACCGAAATGCACCTTAACAATGCCGAAACA 794  
 QY 221 LeuLysAlaGlnThrLysLeuGlnTyrPheAlaValAspArgThrAsnLysLeuLeu 240  
 DB 795 CTGAAAGCCCAAACTGGAATATTTCCAGATGACAGAAACCAATAAAAATTTGCTC 854  
 QY 241 IleLysProLysThrAsnAlaPheGlnAlaIleTyrLysGlnAsnTyrAlaLeuTyrMet 260  
 DB 855 ATCAAAACCAAAACCAATGCGTTTGAACCTGCTTAAAGAAATTTACGCAATGTGGATG 914  
 QY 261 GlyProTyrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer 280  
 DB 915 GGGCGGTTAAAGTAAGCAAAAGAAATCAACCGACGAAAGATTATGTCGATTTCTCC 974  
 QY 281 AspIleAspProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGlnAlaAspAsn 300  
 DB 975 GATATCCGCGCATACGCGCAATCATACGGGTAACTCCGCCCATCGGTAGAGCGTATTAAC 1034  
 QY 301 SerHisGlnGlyTyrGlyTyrSerAspGlnAlaValArgGlnHisAspGlnGlyLeuPro 320  
 DB 1035 AGTCATGAGGGGATGATACGCAATGACGATGACAGCTGCAACAACATGACAGGGCACTT 1094

RESULT 2  
 AR393746 1136 bp DNA linear PAT 18-DEC-2003  
 LOCUS AR393746  
 DEFINITION Sequence 6 from patent US 6617128.  
 ACCESSION AR393746  
 VERSION AR393746.1 GI:40120650  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1136)  
 AUTHORS Meyer,T.F., Rudel,T., Schuenepflug,I., Maier,J., Eickernjager,S.,  
 Schwan,T. and Fischer,B.  
 TITLE Nucleic acid molecules encoding proteins which impart the adhesion  
 of neisseria cells to human cells  
 JOURNAL Patent: US 6617128-A 6 09-SEP-2003;  
 FEATURES Location/Qualifiers  
 source 1..1136  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Alignment Scores: 3.68e-124 Length: 1136  
 Pred. No.: 1645.00 Matches: 320  
 Score: 1645.00 Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 100.00% Gaps: 0

US-10-617-835-4 (1-320) x AR393746 (1-1136)

QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 DB 135 ATGGGGGACGGCGCTGATACCTAATTCCTTTTGGTTTAAATTATCCGCTGGCGG 194  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyIleAspPheAlaValGlnGlnLeu 40  
 DB 195 ACACGACAGGATTCATCCGATGCGGAGGCAAAACGCTTCGGCTGGAACAGAACTT 254  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValIleAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 DB 255 GTGGCCGCTTCGCGAGAGCTGCCGTTAAAGCATGATTTACAGCATTTACACGACGA 314  
 QY 101 AspTyrThrTyrProArgTyrGlyLeuThrAlaGlnThrSerGlyLeuThrGly 80  
 DB 435 GATTACACCTTATCCGCTTACGAAACCAACCGCTGAAACAAATCATGAGCGTTGACGGGT 494  
 QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 DB 495 TTAACTCATCTTTATCTAATCACTTAATGCCCTGACCTCGCGACCACTTACAGACGT 554  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
 DB 555 AGCGGAATAGAGAGAGCTGGGCTTAATATTTGGCGGAGTGGGGATTTATCGAAATGAA 614  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 DB 615 ACCTTGACGACCAACCCCGCGACACTGCTTTCTTCCACTTGGTACAGACCGTATTT 674  
 QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 DB 675 TTCTGCGCGGATAGACGTTGTTTCTCTGCCAATGCCGATACAGATGTTTATTAAAC 734  
 QY 201 IleAspValPheGlyThrIleAsnAsnArgThrGluMetHisLeuTyrAsnAlaGlnThr 220

```

Db      735 ATCGACGATTTCGACACGATACGACAGACCGAATACCTTATACATCCGAAACA 794
Qy      221 LeuYbAaIaGInThrLysLeuGluTyrPheAaVaIaAspArgThrAsnLysLeuLeu 240
Db      795 CTGAAGCCCAACAACTGGATATATTTCGAGTAGACAGAACCAATMAAAATTTGCTC 854
Qy      241 IleLysProLysThrAsnAlaPheGluAlaIaIaTyrLysGluAsnTyrAlaLeuTyrPhe 260
Db      855 ATCAAAACCCAAACCAATGCGTTTGAAGCTGCTTAAAGAAATTTACGATGTGGAG 914
Qy      261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValaAspPheSer 280
Db      915 GGGCGGCTTAAGTAAAGTAAAGCAAGCAATCAACCGAGAGATTAATGCTCATTTCTCC 974
Qy      281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluIaAspAsn 300
Db      975 GATATCCGGCCATACCGCAATCATACGGGTAATCCGCCCATCCGTAGAGGCTGATAC 1034
Qy      301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValaArgGlnHisArgGlnGlyInPro 320
Db      1035 AGTCATGAGGGGTATGATACGATACGATGACGATGACGACATAGACAAAGGCAACT 1094

RESULT 3
LOCUS   A61821 3287 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 1 from Patent WO9711181.
ACCESSION A61821
VERSION A61821.1 GI:3715995
KEYWORDS
SOURCE  Neisseria gonorrhoeae
        Neisseria gonorrhoeae
        Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
        Neisseriaceae; Neisseria.
REFERENCE 1
AUTHORS Meyer,T.F., Rudel,T., Scheuerflug,I., Fischer, Eckhard and
        Maier,J.
TITLE   NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE
        ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
        Patent: WO 9711181-A 1 27-MAR-1997;
        MAX PLANCK GEBELLSCHAFT (DE)
COMMENT  Other publication DE 19534579 970320.
FEATURES
        source
            1..3287
            /organism="Neisseria gonorrhoeae"
            /mol_type="unassigned DNA"
            /strain="MS11"
            /db_xref="taxon:485"
            /clone_lib="GENOMISCHE BIBLIOTHEK IN PBA"

ORIGIN
Alignment Scores:
Pred. No.: 1.38e-123 Length: 3287
Score: 1645.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x A61821 (1-3287)
Qy      1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValaPheIleLeuSerAlaCysGly 20
Db      563 ATGCGGCGACGGCTCTGATACCTATCTTTTCGATTATTTATTTATTCGCCCTCGCGG 642
Qy      21 ThrLeuThrGlyIleProSerHisGlyGlyIleLysArgPheAlaValaGluGlnGluLeu 40
Db      643 ACACGACACGATTCATCCATCGCATGCGGAGGCAACCGTTTCGGGTGGAACAAGAACTT 702
Qy      41 ValaIaAspSerAlaArgAlaAlaValaLysAspMetAspLeuGlnAlaLeuHisGlyArg 60
Db      703 GTGGCGGCTTCTGCGACAGCTGCGTTAAAGACATGATTTTACAGGCAATTACACGAGCA 762

```

```

Qy      61 LysValaIaLeuTyrIleAlaThrMetGlyAspGlnGlySerLeuThrGlyGly 80
Db      763 AAAGTTGATGATTGATATGCAACTATGAGCGACCAAGGCTACAGGAGTTTGACAGGGGCT 822
Qy      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValaArgThr 100
Db      893 CGCTACTCATTTGATGACCTGATTTGGCGGCAATCATTAACAGCCCTCGCGTCCGAC 882
Qy      101 AspTyrThrTyrProArgTyrGlyThrAlaGluThrThrSerGlyGlyLeuThrGly 120
Db      883 GATTACACCTTATCCCGTTACGAAACCAACCGGTGAACACATCAGCGCGTTTACCGGCT 942
Qy      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db      943 TTACCACTTCTTTATCTTACACTTAATGCCCTGACACTCTGCGACCCCAATCAAGCGCT 1002
Qy      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160
Db      1003 AGCGGAAGTAAAGAGAGCTGGGCTTAATAATTGGCGGAGTGGGGATTATCGAAATGAA 1062
Qy      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValaGlnThrValaPhe 180
Db      1063 ACCTTGACGACCAACCCCGGACACCTGCTTTCTTCCCACTTGCTACAGACCGTATTT 1122
Qy      181 PheLeuArgGlyIleAspValaLysProAlaAspThrAspValaPheIleAsn 200
Db      1123 TTCCTCGCGGCAATGACGTTGTTCTCTCGCAATGCGATACAGATGTTATTTATAC 1182
Qy      201 IleAspValaPheGlyThrIleAsnAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220
Db      1183 ATCGACGATTCGGAACGATACGCAACGAAACCGAAATGCACTTATACATCCGAAACA 1242
Qy      221 LeuYbAaIaGInThrLysLeuGluTyrPheAlaValaIaAspArgThrAsnLysLeuLeu 240
Db      1243 CTGAAGCCCAACAACTGGAAATATTTCCGATGACAGAACCAATMAAAATTTGCTC 1302
Qy      241 IleLysProLysThrAsnAlaPheGluAlaIaIaTyrLysGluAsnTyrAlaLeuTyrPhe 260
Db      1303 ATCAAAACCCAAACCAATGCGTTTGAAGCTGCTTAAAGAAATTTAGCATTTGGAGT 1362
Qy      261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValaAspPheSer 280
Db      1363 GGGCGGCTTAAGTAAAGTAAAGCAAGCAATCAACCGAGAGATTAAATGCTGATTTCTCC 1422
Qy      281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValaGluIaAspAsn 300
Db      1423 GATATCCGGCCATACCGCAATCATACGGGTAATCCGCCCATCCGTAGAGGCTGATAC 1482
Qy      301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValaArgGlnHisArgGlnGlyInPro 320
Db      1483 AGTCATGAGGGGTATGATACGATACGATGACGATGACACATAGACAAAGGCAACT 1542

RESULT 4
LOCUS   AR393745 3287 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6617128.
ACCESSION AR393745
VERSION AR393745.1 GI:40120648
KEYWORDS
SOURCE  Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3287)
AUTHORS Meyer,T.F., Rudel,T., Scheuerflug,I., Maier,J., Eickemeyer,J.,
        Schwan,T. and Fischer,E.
TITLE   Nucleic acid molecules encoding proteins which impact the adhesion
        of neisseria cells to human cells
        Patent: US 6617128-A 1 09-SEP-2003;
        Location/Qualifiers
            1..3287
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN

```

**Alignment Scores:**

Pred. No.:	1,38e+123	Length:	3287
Score:	1645.00	Matches:	320
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-617-835-4 (1-320) X AR393745 (1-3287)

OY	1	MetAsgAlaArgLeuLeuLeuLeuLeuLeuPheSerValPheIleLeuSerAlaCysGly	20
Db	563	ATGGCGGCGACGGCTGCTGATACCTATCTTTTCGATTTTATTTATTCGGCTGGGG	642
OY	21	ThrLeuThrGlyIleProSerHisGlyGlyIleArgPheAlaValGluGlnGluLeu	40
Db	643	ACACTGACGAGGATTCATTCGATGGATGGCGAAGCGTTCGGCGGTGCAACAAACTT	702
OY	41	ValAlaAlaSerAlaArgAlaAlaValIleAspMetAspLeuGlnAlaLeuHisGlyArg	60
Db	703	GTGGCGCGCTTCGCCAGACGCTCGTTAAAGCATGATTTTACAGGCATTACCGACGA	762
OY	61	LysValAlaLeuThrIleAlaIleThrMetGlyAspGlnGlySerGlySerLeuThrGlyIle	80
Db	763	AAAGTTGCAATTCATTGACACTATGGCGAGCAAGATTCAGGCGATTGACAGGGAGGT	822
OY	81	ArgTyrSerIleAspAlaLeuLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr	100
Db	823	CGCTACTCCATTGATGCACTGATTCGGCGGCGAATACATAAACGCCCTCGCGTCCGCACC	882
OY	101	AspTyrThrTyrProArgTyrGlyIleThrAlaGluIleThrThSerGlyGlyLeuThrGly	120
Db	883	GATTTCACCTATCCGCGTTACGAAACACCGCTGAAACAAACACAGCGGTTTGAAGGGGT	942
OY	121	LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly	140
Db	943	TTTAAACACTTCTTTATCTACACTTAATGCCCTCGACCTCTGGCGACCCCATAGACGGT	100
OY	141	SerGlySerArgSerSerLeuGlyLeuAsnIleGlyIleMetGlyAspTyrArgAsnGlu	160
Db	1003	AGCGGAAGTAGAGACGATCTGGGCTTAAATATTGGCGGAGATGGGGGATTAACAAATGAA	106
OY	161	ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisIleValGlnThrValPhe	180
Db	1063	ACCTTGACGACCAACCCCGGCACACATGCTTCTTCTCCACTGTGTACGACCGGATTT	112
OY	181	PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn	200
Db	1123	TTCCGTGGCGGATAGACGTTGTTCTCTCTGCAATAGCCGATACAGATGTTTATTAAAC	118
OY	201	IleAspValPheGlyThrIleArgAsnArgThrGluMetHisIleuTyrAsnAlaGluThr	220
Db	1183	ATCGACGATTCGGAAACGATACGCAACAGAACCGAANTCACCCTATACATATGCCGAAACA	124
OY	221	LeuLysAlaGlnThrLysLeuGlnIleTyrPheAlaValAspArgThrAsnLysLysLeu	240
Db	1243	CTGAAGGCCCAACAAACTGGAAATTTTCGAGTAGACAGAAACAATAAAAATTGGCTC	130
OY	241	IleLysProLysThrAsnAlaPheGluAlaAlaTyrLysGlyAsnTyrAlaLeuThrMet	260
Db	1303	ATCAAAACCAAAACCAATGCGTTTGAGACGTGCTTAAAGAAATTACGCAATTGTGGATG	136
OY	261	GlyProTyrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer	280
Db	1363	GGGCGGTATTAAGTAGCAAAAGAAATCAAAACCGAAGGATTAATGTGTCGATTTCTCC	142
OY	281	AspIleAspProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn	300
Db	1423	GATATCCGGCCATTACGGCATATACGGGTAACTCCGCCCATCGTTAAGGCTGATTAAC	148
OY	301	SerHisGlnGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPro	320

Db					
RESULT 5					
AF142582		3294 bp	DNA	linear	BCT 18-MAY-1999
LOCUS					
Neisseria gonorrhoeae maf gene cluster, complete sequence.					
DEFINITION					
AP142582					
VERSION					
AP142582.1	GI:4838552				
KEYWORDS					
SOURCE					
ORGANISM					
Neisseria gonorrhoeae					
Neisseria gonorrhoeae					
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;					
Neisseriaceae; Neisseria.					
1 (bases 1 to 3294)					
Eickenjaeger,S., Moyer,T.F., Fischer,E., Maier,J., Manning,P.A.,					
Rudel,T., Schenepflug,I., Schulz,E. and Schwan,E.T.					
Direct Submission					
Submitted (13-APR-1999) Molekulare Biologie, MPI fuer					
Infektionsbiologie, Mohlbjoustr. 2, Berlin 10117, Germany					
FEATURES					
source					
1..3294					
/organism="Neisseria gonorrhoeae"					
/mol_type="genomic DNA"					
/strain="MS11"					
/db_xref="taxon:485"					
143..457					
/gene="mafI"					
143..457					
/codon_start=1					
/transl_table=1					
/product="MafI"					
/protein_id="AAD31037.1"					
/db_xref="GI:4838553"					
/translation="MRKSRLSRVYKONKLEIFAVGATFARTAAPBDSIVTDCYRRVDV					
LDAEBSHFRINHSTHRAERONHTNGIGNWMAKRRLRKFDGIKPHFPFLYKECBR					
RP"					
590..1552					
/gene="mafA"					
590..1552					
/gene="mafA"					
1/codon_start=1					
/transl_table=1					
/product="adhesin MafA"					
/protein_id="AAD31038.1"					
/db_xref="GI:4838554"					
/translation="MPARLLPILFSVFILISACGLTGIPSHGGGRFAVEOEIVAAVS					
ARAUKMDLQALHGKVALYIAMGOGSGSLSGRGYSIDALIRGEYINSPARTVY					
TTPYEVTAEITSGELTGLTTLSLTUNAPLSRQSDGSRSGLNTIGMGDRNE					
TLTNPRDIAFSLHVOTVFPRLGIDIVASPNADPDVPIIVDGCTIRNTEMLTVNA					
ETLQAQKLETFPAVRINKULLIKPKTNABEAKENYALIMGPYKSKGIKPEGLGM					
VDFSDIRPGNHGTNSASVEADNSHEGYSDVAEDVHRQGDP"					
1592..3121					
/gene="mafB"					
1592..3121					
/gene="mafB"					
1/codon_start=1					
/transl_table=1					
/product="adhesin MafB"					
/protein_id="AAD31039.1"					
/db_xref="GI:4838555"					
/translation="MNLPICKFMWLFAAATLSILOIPISHANGLEDARLRDMOAKHYPP					
GKKYHLFGNARGSVKNRYCAVQTFFDAVTRPILPTIERPGCFEGIIIGYHFSGHGE					
VHSFPHNDHSKTSYDSFGVDGFTIVQLHRTSGEIHPRADYGDPDGGGPREPGARD					
IYSHVIGTSTTKIKINTVPAPPSDRWLKENAGAAGFLSRADAGKLINENDRNKW					
RANMDDIRIGVOGAIVNPPLTGFQGLGVGATDSAVSPVIYAARKTLQGINHNGLNS					
PEADLAASLILODSAFAVKDGINARSQMADAHPVITATAQTALAVALAAGTVMGKKV					
ELNKTQDWNTNYEKFAKPMQTVDGEMAGNKPKPKSTOOHSTADNNIGLAPAY					
VKPPTSISPTGITDIIRMTKSPPEPKSIENGHKAHGKEGGDITIEDYOKMASDLIS					
KQTSIDKLIGLYQTERHRRVRYDINNIIYVLANPKTFIKITMKPKNLGKYTDGEPKDKMG					
N"					

ORIGIN

## Alignment Scores:

Pred. No.: 1,39e-123 Length: 3294  
 Score: 1645.00 Matches: 320  
 Percent Similarity: 100.00% Conservation: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0

US-10-617-835-4 (1-320) x AF142582 (1-3294)

```

QY 1 MetAArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
DB ATGGCGGACCGGCTGCTGATACCTATCTTTTGGTTTATTATTATTCGCCCTGGCGG 649
QY 21 ThrLeuThrGlyIleProSerHisGlyGlyValArgPheAlaValGluGluLeu 40
DB ACACGACAGGATTCATCGATCGCGAGGCAACCGCTCGCGTCCAGAACAACTT 709
QY 41 ValAlaIaSerAlaArgAlaValAlaValysAspMetAspLeuGlnAlaLeuHisGlyArg 60
DB GTGGCCGCTTTCGCGAGCTGCGCTTAAAGACATGATTTACAGGCATTACACGAGCA 769
QY 61 LysValAlaLeuTyrIleAlaThrMetCglyAspGlnGlySerGlySerLeuThrGlyGly 80
DB AAAGTTGCATTGTACATTGCACTATGGCGCACCAAGTTTCAGGCAGTTGACAGGGGT 829
QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
DB CGCTACTCATTCATGACACTGATTCGCGCGAATCATMAACAGCCCTGCCCGCACCC 889
QY 101 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
DB GATTACACCTTATCCGCGTTACGAAACACCGCTGAAACAACTACAGCGGTTGACGGGT 949
QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB TTAAACACCTTCTTATCTTACACTTATATGCGCTGACCTGCGCACCCCAATCAGAGGT 1009
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetCglyAspTyrArgAsnGly 160
DB AGCGAAGTAGAGAGAGTCTGGGCTTAAATATTGGCGGATGGGGGATTTATGAATGAA 1069
QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB ACCTTGACGACCAACCCGCGCACACTGCTTTCTTCCACTTGGTACAGACCGTATTT 1129
QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
DB TTCCGCGCGGCGATAGACGTTGTTCTCTGCCAATGCCGATACAGATGTGTTATTAAAC 1189
QY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220
DB ATCGACGATTCGGAACGATACGCAACAGAACCGAAATGACCTTATACATCCCAACA 1249
QY 221 LeuValAlaGlnThrThrLeuGluGlyTyrPheAlaValAspArgThrAsnLysLeuLeu 240
DB CTGAAGACCCCAACAACTGAAATATTTCCAGTAGACAGAACCAATTAATAAATTTGCTC 1309
QY 241 IleLysProLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTrpMet 260
DB ATCAAAACCAAAACCAATCGTTTGAAGCTGCTTAAAGAAATTAAGCAATGTGGATG 1369
QY 261 GlyProTyrLysValSerLysGlyIleLysProThrGlyGlyLeuMetValAspPheSer 280
DB GGGCGGCTTAAAGTAGAAGAAATCAAAACGAGAGGATTAATGTGCAATTTCTCC 1429
QY 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGlnAlaAspAsn 300
DB GATATCCCGGCGCATAGCGCAATACGCGTAACTCCGCCCATCCGTAGAGGCTGATTAAC 1489
QY 301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyInPro 320
DB AGTCATAGAGGGGTATGATACGAGCATGAAAGACATGACCAATAGACAAAGGCAACCT 1549
  
```

## RESULT 6

A96050 963 bp DNA linear PAT 07-SEP-2000  
 LOCUS Sequence 83 from Patent WO924578.  
 DEFINITION  
 A96050  
 ACCESSION  
 VERSION A96050.1 GI:6779917  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.

REFERENCE  
 1 Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masignani, V.  
 TITLE Neisserial antigens  
 AUTHORS Patient: WO 924578-A 83 20-MAY-1999;  
 JOURNAL PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);  
 CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)

## FEATURES

source 1..963  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

## ORIGIN

Alignment Scores:  
 Pred. No.: 6.33e-124 Length: 963  
 Score: 1641.00 Matches: 319  
 Percent Similarity: 100.00% Conservation: 1  
 Best Local Similarity: 99.69% Mismatches: 0  
 Query Match: 99.76% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x A96050 (1-963)

```

QY 1 MetAArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
DB ATGGCGGACCGGCTGCTGATACCTATCTTTTTCAGTTTATTATTATTCGCCCTGGCGG 60
QY 21 ThrLeuThrGlyIleProSerHisGlyGlyValArgPheAlaValGluGluLeu 40
DB ACACGACAGGATTCATCGATCGCGAGGCAACCGCTCGCGTCCAGAACAACTT 120
QY 41 ValAlaIaSerAlaArgAlaValAlaValysAspMetAspLeuGlnAlaLeuHisGlyArg 60
DB GTGGCCGCTTTCGCGAGCTGCGCTTAAAGACATGATTTACAGGCATTACACGAGCA 180
QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 80
DB TTAAACACCTTCTTATCTTACACTTATATGCGCTGACCTGCGCACCCCAATCAGAGGT 240
QY 181 LysValAlaLeuTyrIleAlaThrMetCglyAspGlnGlySerGlySerLeuThrGlyGly 80
DB AAAGTTGCATTGTACATTGCACTATGGCGCACCAAGTTTCAGGCAGTTGACAGGGGT 240
QY 201 IleArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
DB CGCTACTCATTCATGACACTGATTCGCGCGAATCATMAACAGCCCTGCCCGCACCC 300
QY 241 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
DB GATTACACCTTATCCGCGTTACGAAACACCGCTGAAACAACTACAGCGGTTGACGGGT 360
QY 301 GlnThrAspCCTATCCGCGTTACGAAACACCGCTGAAACAACTACAGCGGTTGACGGGT 360
QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB TTAAACACCTTCTTATCTTACACTTATATGCGCTGACCTGCGCACCCCAATCAGAGGT 420
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetCglyAspTyrArgAsnGly 160
DB AGCGAAGTAGAGAGAGTCTGGGCTTAAATATTGGCGGATGGGGGATTTATGAATGAA 480
QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB ACCTTGACGACCAACCCGCGCACACTGCTTTCTTCCACTTGGTACAGACCGTATTT 540
QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
DB TTCCGCGCGGCGATAGACGTTGTTCTCTGCCAATGCCGATACAGATGTGTTATTAAAC 600
  
```

QY 201 IleaspaValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220  
 DB 601 ATCGACGATTCGGAACGATACGCAACGAAACCGAAATGACCTTATCAATGCCAACA 660  
 QY 221 LeuYsaIaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLeuLeu 240  
 DB 661 CTGAAGCCCAAACTGGATATTTCCGAGTAGACAGAACCAATTAATAATTGCTC 720  
 QY 241 IleLysProLysThrAsnAlaPheGluAlaIleTyrLysGluAsnTyrAlaLeuTyrPhe 260  
 DB 721 ATCAAAACCAAAACCAATGCGTTGAGCTGCTTAAGAAATAATTAAGCAATTTAGCATTTGAGATG 780  
 QY 261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSer 280  
 DB 781 GGGCGGTTAAAGTAGCAAGCAAGCAATCAACCGAAGCAAGATTAATGTCATTTCTCC 840  
 QY 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 DB 841 GATATCCAAACCATACGCAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATAC 900  
 QY 301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGluGlyGlnPro 320  
 DB 901 AGTCATGAGGGGTATGATACAGCGATGAGCAATGACCAATTAAGCAAGGCAACT 960  
 RESULT 7  
 AX043997 963 bp DNA linear PAT 24-NOV-2000  
 LOCUS Sequence 76 from Patent WO0066791.  
 DEFINITION AX043997  
 ACCESSION AX043997.1 GI:11342888  
 VERSION AX043997.1 GI:11342888  
 KEYWORDS  
 SOURCE .  
 ORGANISM Neisseria gonorrhoeae  
 Neisseria gonorrhoeae  
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 Neisseriaceae; Neisseria.  
 REFERENCE  
 1 Pizzo, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,  
 Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarlato, V.,  
 Scarlato, V., Rappuoli, R., Frazer, C. M. and Grandi, G.,  
 Neisseria genomic sequences and methods of their use  
 Patent: WO 0066791-A 76 09-NOV-2000;  
 TITLE CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)  
 JOURNAL  
 FEATURES  
 source 1..963  
 location/Qualifiers  
 1..963  
 /organism="Neisseria gonorrhoeae"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:485"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6,33e-124 Length: 963  
 Score: 1641.00 Matches: 319  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.69% Mismatches: 0  
 Query Match: 99.76% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-617-835-4 (1-320) x AX043997 (1-963)  
 QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 DB 1 ATGCGGGGACGGCTGATACCTATCTTTTTCAGTTTATTTATTCGCGCTGCGG 60  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGlnGluLeu 40  
 DB 61 ACACCTGACAGGATTCATCGCATGCGGAGCAAAACCTTCGCGGTGAAACAAGACTT 120  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 DB 121 GTGCGCGCTTCTGCCACAGCTGCGCTTAAACATGATTTTACAGGCACTTACACGAGCA 180  
 QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80

DB 181 AAAGTGCATTTGATCATTTGCAACTATGGCGCACCAAGGTTACAGGACGAGGGGT 240  
 QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaArgThr 100  
 DB 241 CGCACTCCATTTGATGACCTGATTCGGCGCAATACATTAACAACCCCTGCGTCCGACCC 300  
 QY 101 AspTyrThrTyrProArgTyrGlyThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120  
 DB 301 GATTAACCTTATCCGCGTTACGAACCAACCGCTGAACACATCAAGCGGTTTACGGGT 360  
 QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 DB 361 TTAACCACTTCTTATATCACTTAATGCCCTGCACCTCGCCACCAATTCAGACAGT 420  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
 DB 421 AGCGAAGTAGAAGACAGCTGGGCTTAATAATTTGGCGGATGGGAGATTATCGAAATGAA 480  
 QY 481 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 DB 481 ACCTTGACGACCAACCGCGGACACATGCTTTCTTCCCACTTGTCGACGACGATTT 540  
 QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 DB 541 TTCTCGCGCGCATACGATGCTTTCTCTCCGCAATGCCGATACAGATGTTGTTATTAC 600  
 QY 201 IleaspaValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220  
 DB 601 ATCGACGATTCGGAACGATACGCAACGAAACCGAAATGACCTTATCAATGCGAACA 660  
 QY 221 LeuYsaIaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLeuLeu 240  
 DB 661 CTGAAGCCCAAACTGGATATTTCCGAGTAGACAGAACCAATTAATAATTGCTC 720  
 QY 241 IleLysProLysThrAsnAlaPheGluAlaIleTyrLysGluAsnTyrAlaLeuTyrPhe 260  
 DB 721 ATCAAAACCAAAACCAATGCGTTGAGCTGCTTAAGAAATAATTAAGCAATTTAGCATTTGATG 780  
 QY 261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSer 280  
 DB 781 GGGCGGTTAAAGTAGCAAGCAAGCAATCAACCGAAGCAAGATTAATGTCATTTCTCC 840  
 QY 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 DB 841 GATATCCAAACCATACGCAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATAC 900  
 QY 301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGluGlyGlnPro 320  
 DB 901 AGTCATGAGGGGTATGATACAGCGATGAGCAATGACCAATTAAGCAAGGCAACT 960  
 RESULT 8  
 A96046 963 bp DNA linear PAT 07-SEP-2000  
 LOCUS Sequence 79 from Patent WO924578.  
 DEFINITION A96046  
 ACCESSION A96046  
 VERSION A96046.1 GI:6779915  
 KEYWORDS  
 SOURCE .  
 ORGANISM unidentified  
 unidentified  
 unclassified.  
 REFERENCE  
 1 Pizzo, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masignani, V.  
 Neisseria antigens  
 TITLE PIZZA MARIARAZZA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);  
 CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)  
 JOURNAL  
 FEATURES  
 source 1..963  
 location/Qualifiers  
 1..963  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 ORIGIN

## Alignment Scores:

Pred. No.: 2,34e-123 Length: 963  
 Score: 1634.00 Matches: 317  
 Percent Similarity: 99.69% Conservative: 2  
 Best Local Similarity: 99.06% Mismatches: 1  
 Query Match: 99.33% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x A96046 (1-963)

```

QY 1 MetATGAAATGLeuLeuLeuProLeuPheSerValPheLeuSerAlaCysGly 20
DB 1 ATGCAAGCAGCGCTGCTGATACCTATCTTTTTCAGTTTATTTATTCACCGCTGGGG 60
QY 21 ThrLeuThrglyLeuProSerHISglYgLySargPheAlaValGlnGlnLeu 40
DB 61 ACACGACAGCATTCATCCATCGCATGGCGGAGGTTAAACGCTTTCGGGTGCAACAAGACTT 120
QY 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHISglYarg 60
DB 121 GTGGCCGCTTTCGACAGAGCTCCGTTAAAGACATGATTTTACAGGCATTCACGAGACA 180
QY 61 LysValAlaLeuTyrlleAlaThrmecGlyAspGlnGlySerGlySerLeuThrgly 80
DB 181 AAAGTTGATGTTGATTCATTCGCACTATGGCGCAAGGTTTCAGGCAGTTTGCAGGGGGT 240
QY 81 ArgTySerIleAspAlaLeuIleArgGlyGlyTyrlleAsnSerProAlaValArgThr 100
DB 241 CGCTACTCATTCATGATGCACTGATTCGTGGCGAATCATTAACAGCCCTGCGCTGACC 300
QY 101 AspTyThrTyrlleProArgTyrlleThrAlaGlnThrThrseryGlyLeuThrgly 120
DB 301 GATTACACCTATTCACGTTACGAAACCAACCGCTGAAACAACTCAAGCGGTTTGAACAGT 360
QY 121 LeuThrThrserySerThrseryThrseryThrseryThrseryThrseryThrsery 140
DB 361 TTACCACTCTTTTATCTACACTTAAATGCCCTGCACTCTTCGCAACCAATCAAGAGCT 420
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlymeGlyAspTyArgAnglu 160
DB 421 AGCGAAGATTAAGAGCTGGCTTAAATATTTGGCGGAGTGGGGATTATCGAATGA 480
QY 161 ThrLeuThrThrseryProArgThrseryThrseryThrseryThrseryThrsery 180
DB 481 ACCTTGACGACTAACCCGCGGACGACCTGCTTCTTCCACTTGTCAGACCGCATATT 540
QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrseryPheIleAsn 200
DB 541 TTCTGCGCGGATGAGCTGTGTCTCTCGCAATGCCGATACAGATGTGTTATTAAC 600
QY 201 IleAspValPheGlyThrseryThrseryThrseryThrseryThrseryThrsery 220
DB 601 ATCGACGATTCGCAACGATTCGCAACGAAATGCACTTAACATGATCCGAAACA 660
QY 221 LeuTyValAlaGlnThrseryLeuGlyTyrllePheAlaValAspArgThrseryLeuLeu 240
DB 661 CTGAAGGCCAAACAAACTGGAATATTTCCGACATGACAGAACCAATTAATAATTCGTC 720
QY 241 IleLysProLysThrseryThrseryThrseryThrseryThrseryThrsery 260
DB 721 ATCAACCAACAAACCAATGCCGTTTGAAGCTGCTTAAAGAAATTAACGATTCGATG 780
QY 261 GlyProTyrlleValSerIleGlyIleLysProThrglyGlyLeuMetValAspPheSer 280
DB 781 GGGCGGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 840
QY 281 AspIleArgProTyrlleValSerIleGlyIleLysProThrglyGlyLeuMetValAspPheSer 300
DB 841 GATATTCGACCATTCAGGCAATCATCGGGTAACTCCGCCCATTCGCTGAAGGCTGATTAAC 900
QY 301 SerHISglGlyTyrlleTyrlleSerAspGlnAlaValArgGlnHISArgGlnGlyInPro 320
  
```

DB 901 ACTCATGAGGGTATGATACAGCATGATGAGTGCACAACTATGACAAAGACACT 960

RESULT 9  
 AX043995  
 LOCUS  
 DEFINITION Sequence 74 from Patent WO0066791.  
 ACCESSION AX043995  
 VERSION AX043995.1 GI:11342887  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Neisseria meningitidis  
 Neisseria meningitidis  
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 Neisseriaceae; Neisseria.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 source  
 location/Qualifiers  
 1..963  
 /organism="Neisseria meningitidis"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:487"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,34e-123 Length: 963  
 Score: 1634.00 Matches: 317  
 Percent Similarity: 99.69% Conservative: 2  
 Best Local Similarity: 99.06% Mismatches: 1  
 Query Match: 99.33% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x AX043995 (1-963)

```

QY 1 MetATGAAATGLeuLeuLeuProLeuPheSerValPheLeuSerAlaCysGly 20
DB 1 ATGCAAGCAGCGCTGCTGATACCTATCTTTTTCAGTTTATTTATTCACCGCTGGGG 60
QY 21 ThrLeuThrglyLeuProSerHISglYgLySargPheAlaValGlnGlnLeu 40
DB 61 ACACGACAGCATTCATCCATCGCATGGCGGAGGTTAAACGCTTTCGGGTGCAACAAGACTT 120
QY 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHISglYarg 60
DB 121 GTGGCCGCTTTCGACAGAGCTCCGTTAAAGACATGATTTTACAGGCATTCACGAGACA 180
QY 61 LysValAlaLeuTyrlleAlaThrmecGlyAspGlnGlySerGlySerLeuThrgly 80
DB 181 AAAGTTGATGTTGATTCATTCGCACTATGGCGCAAGGTTTTCAGGCAGTTTGCAGGGGGT 240
QY 81 ArgTySerIleAspAlaLeuIleArgGlyGlyTyrlleAsnSerProAlaValArgThr 100
DB 241 CGCTACTCATTCATGATGCACTGATTCGTGGCGAATCATTAACAGCCCTGCGCTGACC 300
QY 101 AspTyThrTyrlleProArgTyrlleThrAlaGlnThrThrseryGlyLeuThrgly 120
DB 301 GATTACACCTATTCACGTTACGAAACCAACCGCTGAAACAACTCAAGCGGTTTGAACAGT 360
QY 121 LeuThrThrserySerThrseryThrseryThrseryThrseryThrseryThrsery 140
DB 361 TTACCACTCTTTTATCTACACTTAAATGCCCTGCACTCTTCGCAACCAATCAAGAGCT 420
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlymeGlyAspTyArgAnglu 160
DB 421 AGCGAAGATTAAGAGCTGGCTTAAATATTTGGCGGAGTGGGGATTATCGAATGA 480
QY 161 ThrLeuThrThrseryProArgThrseryThrseryThrseryThrseryThrsery 180
DB 481 ACCTTGACGACTAACCCGCGGACGACCTGCTTCTTCCACTTGTCAGACCGCATATT 540
  
```

Qy 181 PheleuArgGlyIleAspValIleSerProAlaIleAspThrAspValPheIleasn 200  
 Db 541 TTCCTGGCCGGCATAGAGCTTCTCTCTGCCAATGCCGATACAGATGTGTTATTAC 600  
 Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyraAsnIleGluThr 220  
 Db 601 ATCGACGATTCCGAACGATACGCAACAGAACCGAAATGCACTTATCAATGCCGAACA 660  
 Qy 221 LeuysAlaGlnThrIleLeuGluThrPheAlaValAspArgThrAsnIleLeuLeu 240  
 Db 661 CTGAAGCCCAACAACTGGAATATTTCCGATGACAGAACCAATPAAAAATTTGCTC 720  
 Qy 241 IleTyProIysThrAsnAlaPheGluIleAlaIleTyArgIleAsnIleLeuThrMet 260  
 Db 721 ATCAACCAAAAAACCAATGCCGTTTGAACCTGCTTAAAGAAATTAAGCATTTGCGATG 780  
 Qy 261 GlyProTyIysValSerIleGlyIleTyProThrGluGlyLeuMetValAspPheSer 280  
 Db 781 GGGCGGTAAAGTAAGCAAGAAAGAAATTAACCGAAGGATTAATGCTGATTTCTCC 840  
 Qy 281 AspIleArgProTyArgIleAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 Db 841 GATATCCGACCATACGCAATCATACGGGTAACTCCGCCATCCGTAGAGGCTGATAC 900  
 Qy 301 SerHisGluGlyTyArgIleTySerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320  
 Db 901 AGTCATGAGGGGTATGATACGATGACGATGATAGTACGACCAACATACCAAGGACCACT 960  
 RESULT 10  
 A96048 A96048 963 bp DNA linear PAT 07-SEP-2000  
 LOCUS Sequence 81 from Patent WO924578.  
 DEFINITION A96048  
 ACCESSION A96048.1 GI:6779916  
 VERSION  
 KEYWORDS  
 ORGANISM  
 SOURCE  
 unidentifed  
 unidentifed  
 unclassified.  
 REFERENCE  
 1  
 AUTHORS  
 TITLE  
 JOURNAL  
 Neisserial antigens  
 Patent: WO 924578-A 81 20-MAY-1999;  
 PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);  
 CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)  
 FEATURES  
 source  
 1. 963  
 Location/Qualifiers  
 /organism="unidentifed"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.51e-122 Length: 963  
 Score: 1624.00 Matches: 315  
 Percent Similarity: 99.69% Conservatave: 4  
 Best Local Similarity: 98.44% Mismatches: 1  
 Query Match: 98.72% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-617-835-4 (1-320) x A96048 (1-963)  
 Qy 1 MeArGAlaArGLeuLeuIleProIleuPheSerValPheIleLeuSerAlaCysGly 20  
 Db 1 ATGCAAGACAGCGCTGCTATACCTATTCTTTTTCAGTTTATTATTATCCGCCGCGGG 60  
 Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyIleTyArgPheAlaValGluGlnGluLeu 40  
 Db 61 ACACGTGACAGGATATCCATCCATGCGGAGGTAAACCTTTGCGGTGGAACAAGAACTT 120  
 Qy 41 ValAlaIleSerAlaArGAlaAlaValIleTyAspMetArPLeuGlnAlaLeuHisGlyArg 60  
 Db 121 GTGGCGCTCTTCGACAGAGTGCCTTAAAGACATGATTTACAGGCATTTACAGGACGA 180

Qy 61 LysValAlaLeuTyriIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
 Db 181 AAAGTTGATTGTATCATTTGCAACTATGGGCGACCAAGGTTCAAGGCAATTTACAGGGGGT 240  
 Qy 81 ArgTySerIleAspAlaLeuIleArgGlyGlyThrIleAsnSerProAlaValArgThr 100  
 Db 241 CGTACTCCATTTGATGACATGATTCGTGGCGAATACATAAACACCCCTGCGCTCGTACC 300  
 Qy 101 AspTyThrTyProArgTyArgIleThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120  
 Db 301 GATTACACCTATCCACGTTACGAAACCAACGCTGAAACAAATACAGGGGTTTACAGGT 360  
 Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 Db 361 TTAAACCTCTTTATATCATCACTTAATGCCCTTCACTCTGCGCACCAATCAGACGGT 420  
 Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyArgAsnGlu 160  
 Db 421 AGCGAAGTAAGAACACACTGGCTTAAATATGGCGGATGGGGATTTATCGAAATGAA 480  
 Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 Db 481 ACCTTGACGACTAACCGCGGACACTGCTTTTCCCATTTGATACAGACCGCATTTT 540  
 Qy 181 PheLeuArgGlyIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleasn 200  
 Db 541 TTCCTGGCCGGCATACAGAGCTTCTCTCTGCCAATCCGATACGATGCTTATTATTAC 600  
 Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyraAsnIleGluThr 220  
 Db 601 ATCCACGTAATTCGAAACGATACGCAACCAAGAAATGCACTTATCAATGCGGAACA 660  
 Qy 221 LeuysAlaGlnThrIleLeuGluThrPheAlaValAspArgThrAsnIleLeuLeu 240  
 Db 661 CTGAAGCCCAACAACTGGAATATTTCCGATGACAGAACCAATPAAAAATTTGCTC 720  
 Qy 241 IleTyProIysThrAsnAlaPheGluIleAlaIleTyArgIleAsnIleLeuThrMet 260  
 Db 721 ATCAACCAAAAAACCAATGCCGTTTGAAGCTGCTTAAAGAAATTAAGCATTTGCGATG 780  
 Qy 261 GlyProTyIysValSerIleGlyIleTyProThrGluGlyLeuMetValAspPheSer 280  
 Db 781 GGACCGGTAAAGTAAGCAAGAAAGAAATTAACCGAAGGATTAAGTGTGATTTCTCC 840  
 Qy 281 AspIleArgProTyArgIleAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 Db 841 GATATCCGACCATACGCAATCATATGGGTAACTCTCCCATCCGATAGAGGCTGATAC 900  
 Qy 301 SerHisGluGlyTyArgIleTySerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320  
 Db 901 AGTCATGAGGGGTATGATACGATGACGATGAGCGACGACATACCAAGGCGCAACT 960  
 RESULT 11  
 AX043999 AX043999 963 bp DNA linear PAT 24-NOV-2000  
 LOCUS Sequence 78 from Patent WO0066791.  
 DEFINITION AX043999  
 ACCESSION AX043999  
 VERSION AX043999.1 GI:11342889  
 KEYWORDS  
 ORGANISM  
 Neisseria meningitidis  
 Neisseria meningitidis  
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 Neisseriaceae; Neisseria.  
 REFERENCE  
 1  
 AUTHORS  
 TITLE  
 JOURNAL  
 Pizza, M., Hickey, E., Peterson, J., Tetrelin, H., Venter, J. C.,  
 Maignani, V., Galeotti, C., Mora, M., Ratti, G., Scarlato, V.,  
 Scarlato, V., Rappuoli, R., Frazzer, C. M. and Grandi, G.  
 Neisseria genomic sequences and methods of their use  
 Patent: WO 0066791-A 78 09-NOV-2000;  
 CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)  
 FEATURES  
 source  
 1. 963  
 Location/Qualifiers  
 /organism="Neisseria meningitidis"





```

US-10-617-835-4 (1-320) x NMA122491 (1-340806)
1 MetatglaalargleuenuillepfoilleleuPheserValpheiilleuSerallacyGgly 20

gene      complement (3561. .3866)
          /label=DUS
          /gene="nuoK"
          complement (3561. .3866)
          /gene="nuoK"
          /EC number="1.6.5.3"
          /note="NMA0005, nuoK, NADH dehydrogenase I chain K, len:
          101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase
          I chain K (EC 1.6.5.3) (102 aa) _fasta scores: E():
          1.1e-24; 65.0% identity in 100 aa overlap. Contains Pfam
          match to entry PF00420 oxidored_q2,
          NADH-ubiquinone/plastocquinone oxidoreductase chain 4L"
          /codon_start=1
          /transl_table=11
          /product="NADH dehydrogenase I chain K"
          /protein_id="CAB83325.1"
          /db_xref="GI:7378783"
          /db_xref="GOA:Q9J0U9"
          /db_xref="UniProt/TrEMBL:Q9J0U9"
          /translation="MTITHTYLVGALPLFGISAMGIPMNKKNVLVLDSTIELMLLVN
          FNFIASFQHLGDPAGQIFVEFVLTVAASASIGLAIMTVLYRRNRQTINVADDELKG"
          complement (3564. .3863)
          /gene="nuoK"
          /note="Pfam match to entry PF00420 oxidored_q2,
          NADH-ubiquinone/plastocquinone oxidoreductase chain 4L,
          score 141.50, E-value 1.5e-38"
          complement (3863. .4534)
          /gene="nuoJ"
          complement (3863. .4534)
          /gene="nuoJ"
          /EC number="1.6.5.3"
          /note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len:
          223 aa; similar to many e.g. NUOJ_RHOCA NADH dehydrogenase
          I chain J (EC 1.6.5.3) (202 aa), _fasta scores: E():
          9.7e-21; 41.3% identity in 201 aa overlap. Contains Pfam
          match to entry PF00499 oxidored_q3,
          NADH-ubiquinone/plastocquinone oxidoreductase chain 6"
          /codon_start=1
          /transl_table=11
          /product="NADH dehydrogenase I chain J"
          /protein_id="CAB83326.1"
          /db_xref="GI:7378784"
          /db_xref="GOA:Q9JX90"
          /db_xref="UniProt/TrEMBL:Q9JX90"
          /translation="MTFQILFTYIPAVIILYXGAKTWTAKNPVHAALHLVLTFCVGM
          LMLMLDEFLFGLTVLVVGVGAWVLFVFMKMLTIEBKRAQFMRRAFPAGVGTLL
          AVALLIILVNPKTDILAAFLGMLKDI PADYNNIRDLGSRITYDYLLPELLAVLLIGLV
          AALALVIRKTVNPKRMDPADQVVRADQGRMLVKEAVKVPOTESAESESDLKR
          EECKGA"
          complement (3874. .3879)
          /gene="nuoJ"
          complement (4034. .4525)
          /gene="nuoJ"
          /note="Pfam match to entry PF00499 oxidored_q3,
          NADH-ubiquinone/plastocquinone oxidoreductase chain 6,
          score 77.90, E-value 2e-19"
          complement (4570. .5277)
          /gene="NMA0007"
          complement (4570. .5277)
          /gene="NMA0007"
          /note="NMA0007, unknown, len: 235 aa"

RBS
misc_feature
gene
CDS

misc_feature
gene
CDS

Alignment Scores:
Pred. No.:      2,24e-119      length:      340806
Score:          1624.00         Matches:      315
Percent Similarity: 99.69%      Conservative: 4
Best Local Similarity: 98.44%      Mismatches:  1
Query Match:     98.72%         Indels:      0
DB:             1              Gaps:        0

```

Db	306931	ATCGAAGCAGCGCTGCTGATACCTATCTTTTCAGTTTATTTATTCACCGCTGGCGG	306872
Qy	21	ThrlLeuthrGly11IproSerH1sg1ygl1y1yarsPheAlaValgl1ng1u1eu	40
Db	306871	ACACTGACAGGATTCATTCGATCGGAGGTAACCGTTTGGCGGTCAACAAGACTT	306812
Qy	41	Val1Ala1a1aSer1Ala1a1aVal1Val1yaspMetAspLeuGln1Ala1euH1sg1yarg	60
Db	306811	GTGGCGGCTTCGCGCAGACCTCGGTTAAAGACATGAGATTTCAGGCATTACACGACCA	306752
Qy	61	LysVal1Ala1eLeuTyrl1eal1arHmetG1yaspGlnG1ySerG1ySerLeuthrGly1y	80
Db	306751	MAAGTTGCACTTGTACCTTCACTTATGCGGACCAAGTTTCAGGCACTTTCAGAGGGGT	306692
Qy	81	Arg1ySer1e1aSp1a1e1u1e1a1e1ygl1y1y1easnSerPro1a1aVal1arg1thr	100
Db	306691	CGCTACTTCATTCATGATCAGTTCGTCGGAATACATTAACAGCCCTCGCTCGTAC	306632
Qy	101	Asp1yTrH1yTrProArg1yTrGluTrH1rH1r1a1aGluTrH1rH1rH1rH1rH1rH1rH1r	120
Db	306631	GATTACACCTATCCACGTTACCAAAACACCGCTGAAACAAATCAGGCGGTTGACAGGT	306572
Qy	121	LeuthrH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1r	140
Db	306571	TTTACCACTTCTTTTATCTACCTTAAATGCGCTGCACTTCGCGACCCAAATCAGAGGT	306512
Qy	141	SerG1ySerArgSerSerLeuG1yLeu1a1n1e1g1y1y1e1y1y1y1y1y1y1y1y1y1y	160
Db	306511	AGCGGAGTAAGACGCTGGCTTAATATTTGGCGGATGAGGATTAATGCAATGAA	306452
Qy	161	ThrLeuthrH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1r	180
Db	306451	ACCTTGACGACATACCCCGGACACATGCTTCTTCCACTGTTACAGACCGATATT	306392
Qy	181	PheLeuArgG1y11e1aSp1a1a1e1y1y1e1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y	200
Db	306391	TTTCTGCGGCGCATAGACCTTGTCTTCTGCGCAATGCGGATAGGATGTTTATTAC	306332
Qy	201	11e1aSp1a1a1e1y1y1e1y	220
Db	306331	ATCGACGATTCGGAACGATACGCAACAGAACGAAATGCACTTATACATGCCGAAACA	306272
Qy	221	Leu1yVal1aGln1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1r	240
Db	306271	CTGAAAGCCCAACAAACATGGAATAATTCGACAGTACAGAACCAATTAATAATTTGCTC	306212
Qy	241	11e1yPro1yLeuThrH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1rH1r	260
Db	306211	ATCAAAACCAAAACCAATGCTTGAAGCTGCTTATTAAGAAATTAACGATTTGATG	306152
Qy	261	G1yPro1y	280
Db	306151	GGACCGTATTAAGTAAGCAAGATTAACCGACGAAAGATTAAGTTCGATTTCTCC	306092
Qy	281	Asp1yLeuArgPro1yG1y	300
Db	306091	GATATCCCAACCAATGCGCATATATGAGTAATCTGCGCCATCCGTAAAGGCTATTAAC	306032
Qy	301	SerH1aGlnG1yTrG1yTrSer1aSp1a1a1aVal1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a	320
Db	306031	AGTCATGAGGGGTAATGATACAGCATGAGACAGTCCGACGACATAGACAGGCACTT	305972

## RESULT 13

AE002559

10869 bp DNA linear BCT 25-MAY-2000

Neisseria meningitidis serogroup B strain MCS8 section 201 of 206

of the complete genome.

ACCESSION AE002559 AE002098

VERSION AE002559.2 GI:7413462

KEYWORDS

SOURCE Neisseria meningitidis MCS8

ORGANISM Neisseria meningitidis MCS8

REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.
AUTHORS	1 (bases 1 to 10869) Tetzelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Petersen, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clifton, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H., Qin, H., Vamathavan, J., Gill, J., Scarlato, V., Maignani, V., Pizzi, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R. and Venter, J.C.
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8 (5459), 1809-1815 (2000)
JOURNAL	Science 287 (5459), 1809-1815 (2000)
MEDLINE	20175755
PUBMED	10710307
REFERENCE	2 (bases 1 to 10869) Tetzelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Petersen, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clifton, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H., Qin, H., Vamathavan, J., Gill, J., Scarlato, V., Maignani, V., Pizzi, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R. and Venter, J.C.
AUTHORS	Direct Submission Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA On Apr 4, 2000 this sequence version replaced gi:7227357.
COMMENT	Location/Qualifiers
FEATURES	1..10869 /organism="Neisseria meningitidis MCS8" /mol_type="Genomic DNA" /strain="MCS8" /db_xref="taxon:122586" /note="serogroup: B" 150..524 /gene="NMB2095" 150..524 /gene="NMB2095" /note="similar to PID:1817528 percent identity: 53.99; identified by sequence similarity; putative" /codon_start=1 /transl_table=1 /product="adhesin complex protein, putative" /protein_id="AAF42412.1" /db_xref="GI:7227358" /translation="MKLTITTAISATRLSMAAAGDNPVAKTVSYVCOGCKYKVITGFNKGGLTTASAVINCKRVQMPVNLKSDNVEFTYKEGGYVLTGVMDSKYRKQIPMTAPNQIVFKDCSPR" complement (712..2178) /gene="NMB2096" complement (712..2178) /gene="NMB2096" /note="similar to SP:P33940 PID:424051 GB:U00096 percent identity: 69.73; identified by sequence similarity; putative" /codon_start=1 /transl_table=1 /product="malate:quinone oxidoreductase" /protein_id="AAF42413.1" /db_xref="GI:7227359" /translation="MAEATDVVLVGGIMSATLGVLLKELEPSEWITLIRLEDALESSNANNAAGTSHALCELVNAPLCAANGIIPDARALNIAEQPHVSRQFATVVAEGKLEDSNFTAPVHMSLVNEDHSGYLQKRDAPFTOKLFEMSESTDRNKISDPAVPMRGKDENOVAANASAGETVDVDFGLTRVQWYVLOGGVTEFRHVEDIKRESDGAVLKTADTRNDPGQTLTRPLFLGAGGALITLLOKSGIPEKQVCGRPVSGLPFRNSNPTAEQNAKVTGQASVGAPPWSVPHLDTRVDDKRLHMFQVPAVGRSNPLKQGLNDPLSIHMDNLYPMLCAGWAMPVLYKYLGLGRKESRFSALSYEPAANDDWELITAGQVJOIIXKDSKGVLOFCTEIVAHADSGLAALLCASPGASVAVPLMLRLMQCPPERAPSWDRLEKELVPGYGIKLNENPERADEIIATYAKLDI"

gene	complement (2230. .2319) /gene="NMB2097" complement (2230. .2319) /gene="NMB2097" /note="hypothetical protein; identified by Glimmer2; putative"
CDS	/codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAFA42414.1" /db_xref="GI:7227360" /translation="MLATHSNPEYATYSDTHCRFSLLRFRL" complement (2715. .3305) /gene="NMB2098" complement (2715. .3305) /gene="NMB2098" /note="conserved hypothetical protein; identified by Glimmer2; putative"
gene	/codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAFA42415.1" /db_xref="GI:7227361" /translation="MSKKILPLVSLSGVRLEPLDHHETGLRAVCDGGEVKKGVTS APHPROVDITATLRLAFVAVDDEGRVGTATYHPEPQIPRLDIGTWTYASAA RRTNRITCKKMLDSAPDVLCRCVGRTOIIMLASGALRIERGAEDGVLRRHMLR KDGSVRDVIYVSMLEEDWCNKEIITGLAGYGV" complement (3366. .3947) /gene="NMB2099" complement (3366. .3947) /gene="NMB2099"
CDS	/note="conserved hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAFA42416.1" /db_xref="GI:7227362" /translation="MRNEKRALRRELNGRRSSQMGDVRPAATVKINHLIKRYIKKGR KIGVYPMQKEKRLDGFVRAOAKGCAELYLPIEBSRHMFTPEPADGVQEKRRR AKLHVPPFGKRRKHVDLNLILPVVGMDRLGRLGQAGGYDATLSMKYRLDAKTV GVGPAQCLVDRIPVPAHRSRDLGPFSEGLICF" 4025. .4258 /gene="NMB2100" 4025. .4258 /gene="NMB2100"
gene	4025. .4258 /gene="NMB2100" /note="hypothetical protein; identified by Glimmer2; putative"
CDS	/codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAFA42417.1" /db_xref="GI:7227363" /translation="MMHIVAGYLFVAMTGAAPSTAKRLIYLVFAVLPTVTVT ITVRRNHLHQOEAESEQAPQKDSGTP" 4431. .5159 /gene="NMB2101" 4431. .5159 /gene="NMB2101"
gene	4431. .5159 /gene="NMB2101" /note="similar to SP:P34831 GB:X53651 PID:296030 percent identity: 78.02; identified by sequence similarity; putative"
CDS	/codon_start=1 /transl_table=11 /product="30S ribosomal protein S2" /protein_id="AAFA42418.1" /db_xref="GI:7227364" /translation="MSQITRMQKLEAGVHFGHTRFMNPKAAQYIFGARNKIHLVNL KTLPMFOAKGEVRLVANKGVTLFVGTRQARQDIIEAATRAQMPVDVRYLGMKLT NKTATYQIKRIEELKTALENAAESGFSEKIELMQRDVEKLESLGIDKMGKLPDA IFVIDTGOKGTLVAEKLGIPTVAVDTNNSPGVKYVIFGNDSDSAKAIIRLYRGIA DAVEGKQALQETVAAQEAALAE" 5288. .6142 /gene="NMB2102"

```

CDS
/gene="NM2102"
5288..6142
/gene="NM2102"
/note="similar to GB:D26562.SP:P02997 GB:DI13334 GB:V00343
PID:1208943 percent identity: 69.01; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="elongation factor TS (EF-TS)"
/protein_id="AAFA2419.1"
/db_xref="GI:7227365"
/transl_table="MARIETAKVADLRRAATGLGMECKKALVEABGNPDFKAEILRIIR
SGAAGKLAGRTAEBGLYAINGNVGLVEVNETPVPANDAFVEBANSVATTAAB
KPSSEVELSELVEARRKALIIAKLGEMNSVRRFOVITDANQVYHGALATGEVLYE
KKGSDVARKIGHIIVAKPCVSEAEVDATVEKERRIYTEQAIASGKPDIDIAKRW
EGARRKRLAEITLTNGQAFVNNPDQTVAFSKEKCTEIVISFVRKVGDIGIEKKADVAA
EAAAKAV"
6359..7078
/gene="NM2103"
6359..7078
/gene="NM2103"
/note="similar to GB:D26562.SP:P29464 GB:DI13334 GB:X78809
PID:1208944 percent identity: 73.16; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="uridylylate kinase"
/protein_id="AAFA2420.1"
/db_xref="GI:7227366"
/transl_table="MTQOIKYKRVLLKLSGESLMSGDPGINHDTIVQTGELAEVYK
MGVQVGVIVGSGNIFRQVSAQSGMDAPVTDYKMSMATVNNALAKAPETLGIKAV
QSALSMQOIAETVAPRAIOYLEBGRKATIPAGGNGPFTTIDTAALRGABMNDVIL
KATINDVYIADPKDPSATRYETITPDEALKNLKVMDATAFALCRERKLNTVVFQI
AKESGLKRVITGEBEGTLVHC"
7262..8225
/gene="NM2104"
/note="similar to GP:4838554 percent identity: 99.38;
identified by sequence similarity; putative"
8265..9716
/gene="NM2105"
8265..9716
/gene="NM2105"
/note="similar to GP:4838555 percent identity: 0.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1

```

[illegible]

```

QY      80  YARGTYRSEITLLEAPALALEUILLIARGGLVGLUTYRILLEANSERPROALAVALARGTH 100
DB      7502 TCGTACACCCATGATGACGATTCGGCCAAATACAAACACCCCTGCGCTGAC 7561
QY      100  RASPTYRTHYRPROARGYRGLUTHRTHRALGLUTHRTHRSEGLYGLYLEUTHG1 120
DB      7552 CGATTACACCTATCCAGCTTACGAAACCCGCTAAACACATCAGCGCGTTTACAGG 7621
QY      120  YLEUTHRTHRSEITLLEAPALALEUILLIARGGLVGLUTYRILLEANSERPROALAVALARGTH 140
DB      7622 TTTAACCACTTCTTATCTACACTTAATGCCCTCTGACCTCTCCACCAATCAGACGG 7681
QY      140  YSERGLYSEARSGSERSEITLLEUANSILLEGYGLYMETGLYAPPTYRARGANG1 160
DB      7682 TAGCGAAGTAAACACGCTGGGCTTAATAATTTGGCGGATGGGGATTCGAAATGA 7741
QY      160  UTHRLEUTHRTHRANPPOARGAPRTHRALPHELEUSERHISLEUVALGINTHVALPH 180
DB      7742 AACCTTGACGATTAACCCGCGGACACCTGCTTTTCCCACTTGTRACAGACGCTAT 7801
QY      180  EPHELEUARGGLYLEAPVALVALSERPROALAANSALAEPTHRAPVALPHEILEAS 200
DB      7802 TTTCTGCGCGGATAGACGTTGTTCTCTCCATGCGCATACAGATGTTATTAA 7861
QY      200  NILEAPVALPHEGLYTHRILEARGAANARGTHRGILWETHISLEUYRANALAGLU 220
DB      7862 CATCAACCTATTCGAGACGATACGACGAAACCCAAATGCACTATCAATCAAGCCGAAC 7921
QY      220  FLEUVALAGINTHRYSLEUGLUYRPHALVALASPARGTHRANLVSLEULE 240
DB      7922 ACTGAAGCCCAACAAACTGGAATATTTCCGATGACAGAACCAATMAAAATTTGCT 7981
QY      240  NILEYPROLYSTHRANALAPHEGLUALAALATYRYSGLUANSYRVALALEUTIME 260
DB      7982 CATCAACCAAAACCAATGCGTTTGAAGCTGCTATTAAGAAATTTACGATGTGANT 8041
QY      260  EGLYPROLYSTHRANALAPHEGLUALAALATYRYSGLUANSYRVALALEUTIME 280
DB      8042 GGGGCGGATTAAGATGACGAAAGAAATTAACCGAAGGATTAATGATTTCTC 8101
QY      280  RASPTYRTHYRPROARGYRGLUTHRTHRALGLUTHRTHRSEGLYGLYLEUTHG1 300
DB      8102 CGATTACACCTATCCAGCTTACGAAACCCGCTAAACACATCAGCGCGTTTACAGG 8161
QY      300  NSEHISGLUGLYYRGLYRYSERAPGLUALAVALARGGLINHSARGGLINGLYG1NR 320
DB      8162 CAGTCATGAGGGGTATGATACAGCATGAATGAGCAACATACAGACGAAACC 8221
QY      320  O 320
DB      8222 T 8222

RESULT 14
LOCUS   AX044035 172325 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 114 from Patent WO066791.
ACCESSION AX044035
VERSION   AX044035.1 GI:11342919
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis
          Neisseria meningitidis
          Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
          Neisseriaceae; Neisseria.
REFERENCE
AUTHORS 1 Pizzo, M., Hickey, E., Peterson, J., Tetteijn, H., Venter, J. C.,
          Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarfelli, M.,
          Scarfello, V., Rappuoli, R., Frazer, C. M. and Grandi, G.
          Neisseria genomic sequences and methods of their use
          Patent: WO 066791-A 114 09-NOV-2000;
          CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
SOURCE   1. 172325
          /organism="Neisseria meningitidis"

```

```

/mol_type="unasigned DNA"
/db_xref="taxon:487"
/note="Sequence too long, cut in 8 pieces--seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 1200001 to 1549980
349980 bases--seq 112: 1500001 to 1849980 349980 bases--seq
113: 1800001 to 2149980 349980 bases--seq 114: 2100001 to
227325 172325 bases"

ORIGIN
Alignment Scores:
Pred. No.: 1,68e-119 Length: 172325
Score: 1621.00 Matches: 317
Percent Similarity: 99.38% Conservative: 2
Best Local Similarity: 98.75% Mismatch: 1
Query Match: 98.54% Indel: 1
DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x AX044035 (1-172325)
QY      1  MCAARGALARGLEULEUILLIAPROILLEUPHESERVALPHEILLEUSERALACYSGLY 20
DB      121692 ATGCAAGACGCGCTGCTGATACCTATCTTTTTCAGTTTATTATTATCCGCTGCGGG 121751
QY      21  THRLEUTHRGLYILEPROSERHISGLYGLYLYAARGPHEALVALGLUGLULEU 40
DB      121752 ACACGTGACAGGATTCCTCCATGCGAGGTAAACCTTTGCGGTGAACAAACCTT 121811
QY      41  VALAIALASERVALARGALALAVALLYSAPMECAPLEUGNALALEUHSGLYARG 60
DB      121812 GTGGCGGCTTCTGCGAGGCTGCGCTTAAAGCATGATTTTACAGCATTTACAGGACGA 121871
QY      61  LYSVALAIALEUYRILEALATHMERGLYAPGLINGLYSERGLYSELEUTHR-GLYG1 80
DB      121872 AAAGTTGATTTGATTCATTCACCATGAGCGGCAAGGTTTCAAGGATTTGACAGGGGG 121931
QY      80  YARGTYRSEITLLEAPALALEUILLIARGGLVGLUTYRILLEANSERPROALAVALARGTH 100
DB      121932 TCGTACACCCATGATGACGATTCGGCCAAATACAAACCCCTGCGCTGAC 121991
QY      100  RASPTYRTHYRPROARGYRGLUTHRTHRALGLUTHRTHRSEGLYGLYLEUTHG1 120
DB      121992 CGATTACACCTATCCAGCTTACGAAACCCGCTAAACACATCAGCGCGTTTACAGG 122051
QY      120  YLEUTHRTHRSEITLLEAPALALEUILLIARGGLVGLUTYRILLEANSERPROALAVALARGTH 140
DB      122052 TTTAACCACTTCTTATCTACACTTAATGCCCTCTGACCTCTCGCAACCAATCAGACGG 122111
QY      140  YSERGLYSEARSGSERSEITLLEUANSILLEGYGLYMETGLYAPPTYRARGANG1 160
DB      122112 TAGCGAAGTAAACACGCTGGGCTTAATAATTTGGCGGATGGGGATTTTCGAAATGA 122171
QY      160  UTHRLEUTHRTHRANPPOARGAPRTHRALPHELEUSERHISLEUVALGINTHVALPH 180
DB      122172 AACCTTGACGATTAACCCGCGGACACCTGCTTTTCCCACTTGTRACAGACGCTAT 122231
QY      180  EPHELEUARGGLYLEAPVALVALSERPROALAANSALAEPTHRAPVALPHEILEAS 200
DB      122232 TTTCTGCGCGGATAGACGTTGTTCTCTGCGCAATCCGATACAGATGTTATTAA 122291
QY      200  NILEAPVALPHEGLYTHRILEARGAANARGTHRGILWETHISLEUYRANALAGLU 220
DB      122292 CATCAACCTATTCGAGACGATACGACGAAACCCAAATGCACTATCAATGCGGAAC 122351
QY      220  FLEUVALAGINTHRYSLEUGLUYRPHALVALASPARGTHRANLVSLEULE 240
DB      122352 ACTGAAGCCCAACAAACTGGAATATTTGCGATGACAGAACCAATMAAAATTTGCT 122411
QY      240  NILEYPROLYSTHRANALAPHEGLUALAALATYRYSGLUANSYRVALALEUTIME 260
DB      122412 CATCAACCAAAACCAATGCGTTGAGGCTGCTATTAAGAAATTTACGATTTGANT 122471

```

QY 260 tGlyProTyrIysValSerIysGlyIleuYsProThrGluGlyLeuMetValAspPheSe 280  
Db 122472 GGGGGCGGTAATAGTAAAGCAAGAAATTAACCGACGAAAGATTAAGTCGATTTCTC 122531  
QY 280 TAspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAs 300  
Db 122532 CGATATCCGACCATATCGGCAATCATACGGGTAACTCCGCCCATCTCGTAGAGGTGATTA 122591  
QY 300 nSerIleGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPr 320  
Db 122592 CAGTATAGAGGGGTATGATTAACGCGATGAGTAACTGACCAACATAGACAAGACAACC 122651  
QY 320 o 320  
Db 122652 T 122652

RESULT 15  
A96284  
LOCUS A96284 942 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 317 from Patent WO924578.  
ACCESSION A96284  
VERSION A96284.1 GI:6780034  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Pizsa,M., Scariato,V., Rappuoli,R., Grandi,G. and Maignani,V.  
TITLE Neisserial antigens  
JOURNAL Patent: WO 924578-A 317 20-MAY-1999;  
PIZZA MARIAGRAZIA (IT); SCARILATO VINCENZO (IT); RAPPUOLI RINO (IT);  
CHIRON SPA (IT); GRANDI GUIDO (IT); MALIGNANI VEGA (IT)  
FEATURES  
source location/Qualifiers  
1..942  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.19e-74 Length: 942  
Score: 1030.00 Matches: 197  
Percent Similarity: 78.66% Conservative: 50  
Best Local Similarity: 62.74% Mismatches: 61  
Query Match: 62.61% Indels: 6  
DB: 6 Gaps: 2

US-10-617-835-4 (1-320) x A96284 (1-942)

QY 5 LeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCyGlyThrIleuThrGly 24  
Db 16 CTCCTCATCCCCCTC-----GTACTCACCGCTCGGCGACACTGACCCGCG 60  
QY 25 IleProSerHisGlyGlyIleValysArgPheAlaValGluGlnGlyLeuValAlaAsp 44  
Db 61 ATACCCGCGCCACGGCGCGGCAAAAGCTTTGCGCGTCCGAAACAGAACTGTCGCGCGCATCG 120  
QY 45 AlaArgAlaAlaValIysAspMetAspLeuGlnAlaLeuHisGlyValArgValAlaLeu 64  
Db 121 TCCCGCGCGCGCTCAAAAGAAATGGAATTGTCGCGCTGAAGAGACGCAAAACCGCCCTT 180  
QY 65 TyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyIleArgTyrSerIle 84  
Db 181 TACGCTCTCGTTATGGCGGACCAAGTTCCGGCAACATAGAGGCGGACGCTACTCCATC 240  
QY 85 AspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThrAspTyrThrTyr 104  
Db 241 GACGCACTGATACGGCGCGGCTACCAACAACCCGACAGCGCCGATACAGCTAC 300  
QY 105 ProArgTyrGluThrThrAlaGluThrThrSerGlyIleuThrGlyLeuThrThrSer 124  
Db 301 CCCGCTATGAACTACCGCACCAAAATCCGACGCGCTCTCCGGCGTAAACCACTTCC 360

QY 125 LeuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerArg 144  
Db 361 ACATCGCTTTTGAACGCCCCCGCGCGCCCTGTACGAAACAAACAGACGAAAGCGCA 420  
QY 145 SerSerLeuGlyLeuAsnIleGlyIleMetGlyAspTyrArgAsnGluThrIleuThr 164  
Db 421 CGCTCCCGCGGACGTCCGTCMAAGCGACGGGCGCATACCGCAACGAAACCTTCCTCGCC 480  
QY 165 AspProArgAspThrAlaPheLeuSerHisIleuValGlnThrValPhePheLeuArgGly 184  
Db 481 AACCCCGCGACGTTTCTCTTCAACCACTCATCCAAACGTCCTTCTACTCGCGCGC 540  
QY 185 IleAspValAlaSerProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPhe 204  
Db 541 ATCGAAGTCGTAACCGCCCGAATACCGACACGACGACGATTCGTAACCGTGAAGTTC 600  
QY 205 GlyThrIleArgAsnArgThrGlnMetHisIleuTyrAsnAlaGluThrLeuValGln 224  
Db 601 GGCACCGTCCGCGACCGCTACCGAATCTCACCTCTACAAACCGCGAAACCTTAAAGCCCA 660  
QY 225 ThrIleuGluGluThrPheAlaValAspArgThrAsnIleValIleuLeuIleValProlys 244  
Db 661 ACCAAGCTCGAATTTTCCGCGTGCAGCGGACGACGCGAAGACTGTGATTTGCCCTTAA 720  
QY 245 ThrAsnAlaPheGluAlaAlaTyrIleGluAsnTyrAlaLeuThrMetGlyProTyrIlys 264  
Db 721 ACCGCGCGCTACGAAATCCCAATACCAAGAACAAATACCCCTCGATGGGACCTTACAGC 780  
QY 265 ValSerIysGlyIleValProThrGlnGlyIleuMetValAspPheSerAspIleArgPro 284  
Db 781 GTCGCAAAACCGCAAGGCTCAGACCGCTGATGTCATTTCTCCGACATCAACCC 840  
QY 285 TyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerHisGlyGly 304  
Db 841 TACGCGCACACACCGCCCAAAACCGTCCGACTTCAAAACAAACACGCT--AAAAC 897  
QY 305 TyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGly 318  
Db 898 CCCGATGTCGCAACGAAATCATCCGCGCGCAAGAGGA 939

Search completed: August 19, 2005, 00:57:11  
Job time : 4081 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 18, 2005, 22:56:54 ; Search time 534 Seconds  
(without alignment)  
3547.411 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645  
Sequence: 1 MRARLLPLPSVFLSACG.....SHEGYSDAVRQRGQGP 320

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame.p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US10617835/runcat.18082005.115607.8354/app.query.faeta.1.519  
-DB=N Geneseg.16Dec04 -QFMT=fasta2p -SUPFIX=sp2n.rng -MINMATCH=0.1 -LOOPCL=0  
-IOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biolum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10617835@cgn.1.1.708 @runcat.18082005.115607.8354 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=30 -THREADS=1 -DSPBLOCK=100 -LONOLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N\_Geneseg.16Dec04:\*  
2: genesegn1980s:\*  
3: genesegn1990s:\*  
4: genesegn2000s:\*  
5: genesegn2001s:\*  
6: genesegn2002s:\*  
7: genesegn2003s:\*  
8: genesegn2003bs:\*  
9: genesegn2003cs:\*  
10: genesegn2003ds:\*  
11: genesegn2003fs:\*  
12: genesegn2004as:\*  
13: genesegn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1645	100.0	3287	2	AAT70386 Neisseria
2	1641	99.8	960	10	AB237917 N. gonorr
3	1641	99.8	963	2	AA212008 Neisseria
4	1641	99.8	963	3	AA254614 Neisseria
5	1641	99.8	963	3	AA253688 Neisseria

6	1641	99.8	963	3	AA21297	AA21297 N. gonorr
7	1641	99.8	963	3	AA21582	AA21582 N. gonorr
8	1634	99.3	963	2	AA212006	AA212006 Neisseria
9	1634	99.3	963	3	AA253689	AA253689 Neisseria
10	1634	99.3	963	3	AA254613	AA254613 Neisseria
11	1634	99.3	963	3	AA21296	AA21296 N. mening
12	1634	99.3	963	3	AA21581	AA21581 N. mening
13	1624	98.7	963	2	AA212007	AA212007 Neisseria
14	1624	98.7	963	3	AA254615	AA254615 Neisseria
15	1624	98.7	963	3	AA253690	AA253690 Neisseria
16	1624	98.7	963	3	AA21298	AA21298 N. mening
17	1624	98.7	963	3	AA21583	AA21583 N. mening
18	1621	98.5	92934	3	AA21473	AA21473 N. mening
19	1621	98.5	110000	3	AA21489_7	Continuation (8 of
20	1621	98.5	172325	3	AA21613	AA21613 Neisseria
21	1318	80.1	789	10	AB240799	AB240799 N. gonorr
22	1030	62.6	939	10	AB239696	AB239696 N. gonorr
23	1030	62.6	939	10	AB241775	AB241775 N. gonorr
24	1030	62.6	942	2	AA212111	AA212111 Neisseria
25	1029	62.6	640	2	AA212005	AA212005 Neisseria
26	1029	62.6	640	2	AA21330	AA21330 N. mening
27	1029	62.6	942	2	AA212109	AA212109 Neisseria
28	1029	62.6	110000	3	AA21490_03	Continuation (4 of
29	1029	62.6	110000	3	AA21490_06	Continuation (7 of
30	1029	62.6	110000	3	AA21489_0	AA21489 N. mening
31	1029	62.6	349980	3	AA21607	AA21607 Neisseria
32	1029	62.6	349980	3	AA21608	AA21608 Neisseria
33	1022	62.1	942	2	AA212110	AA212110 Neisseria
34	882.5	53.6	882	10	AA239704	AA239704 N. gonorr
35	789.5	48.0	25509	3	AA21467	AA21467 N. mening
36	650	39.5	591	2	AA212108	AA212108 Neisseria
37	650	39.5	591	3	AA21359	AA21359 N. mening
38	273	16.6	599	3	AA21265	AA21265 N. mening
39	117.5	7.1	4276	4	AB117383	AB117383 Drosophila
40	117.5	7.1	4200	4	AB117307	AB117307 Drosophila
41	106.5	6.5	573	10	AD299982	AD299982 Bacteriophage
42	106.5	6.5	1521	11	ACH99804	ACH99804 Klebsiella
43	105.5	6.4	2196	8	ADA69363	ADA69363 Rice gene
44	104.5	6.4	110000	6	AB269245_01	Continuation (2 of
45	104.5	6.4	110000	6	AB267197_00	AB267197 listeria

## ALIGNMENTS

RESULT 1	
AA270386	standard; DNA; 3287 BP.
XX	
AC	AA270386;
XX	
DT	08-DEC-1997 (first entry)
XX	
DE	Neisseria adhesion proteins.
XX	
KW	Neisseria gonorrhoea; adhesion; lipoprotein; OrfA; Orf1; OrfB; BB.
XX	
OS	Neisseria gonorrhoea.
XX	
FH	
FT	Key
FT	-35_signal
FT	/tag= a
FT	78.83 b
FT	-10_signal
FT	/tag= b
FT	124.128
FT	RBS
FT	/tag= c
FT	136.449
FT	CDS
FT	/tag= d
FT	/label= OrfA
FT	514.519
FT	-35_signal
FT	/tag= e
FT	-10_signal
FT	534.539
FT	/tag= f
FT	571.577
FT	RBS

```

FT      CDS      /*tag= 9
FT      CDS      583..1545
FT      CDS      /*tag= h
FT      CDS      /label= Orf1
FT      RBS      1573..1579
FT      RBS      /*tag= 1
FT      CDS      1585..3114
FT      CDS      /*tag= }
FT      CDS      /label= OrfB
XX      DE19534579-A1.
XX      PD      20-MAR-1997.
XX      PF      18-SEP-1995; 95DE-01034579.
XX      PR      18-SEP-1995; 95DE-01034579.
XX      PA      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX      DR      WPI, 1997-180942/17.
XX      DR      P-PSDB; AAM18784, AAM18785, AAM18786.
XX      PT      Nucleic acids encoding Neisseria adhesion proteins - for therapeutic and
XX      PS      diagnostic use.
XX      PS      Claim 1; Page 11-13; 20pp; German.
XX      CC      OrfA and OrfB in complexes with the protein PilC are capable of adhering
XX      CC      to human cells. Products obtained from the DNA are useful in medicaments,
XX      CC      diagnostic compns. and vaccines, esp. for treatment of Neisseria
XX      CC      gonorrhoea and N. meningitidis infections
XX      SQ      Sequence 3287 BP; 1016 A; 741 C; 752 G; 778 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,65e-158 Length: 3287
Score: 1645.00 Matches: 320
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-617-835-4 (1-320) x AAT70386 (1-3287)
QY      1 MetArgAlaArgLeuLeuIleproIleuPheSerValPheIleLeuSerAlaCysGly 20
DB      583 ATGCGGCGACGCGTCTGATACCTATCTTTTCGTTTATTATTCGCCCTGCGGG 642
QY      21 ThrLeuThrGlyIleProSerHisGlyGlyIysArgPheAlaValGlnGlnIleu 40
DB      643 ACACGACAGGATTCATCGATGCGGAGCGCAACCTTCGCGGTGAAACAAGACTT 702
QY      41 ValAlaAlaSerAlaArgAlaAlaValIysAspMetAspLeuGlnAlaLeuHisGlyArg 60
DB      703 GTGGCGCGCTCTGCGACAGCTGCCGTTAAAGCAATGATTAAACGCGATTACACCGAAGA 762
QY      61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
DB      763 AAAGTGCATTGTACATTGCAACTATGCGCAGCAAGGTTCCAGGCAAGTTTGCAGAGGGGT 822
QY      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
DB      823 CGCTACTCCATTGATGACACTGATTCGCGCGCAATACATTAACAGCCCTGCGTCCGACAC 882
QY      101 AspTyrThrTyrProArgTyrGlyuThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
DB      883 GATTACACCTTATCCCGTTACAAACCAACCGCTGAACCAACATCAGCGGTTTGAACGGGT 942
QY      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB      943 TTAAACCACTCTTTATCTTACACTTAATGCCCCCTGCACTCTCGCGCACCAACATCAGACGGT 1002

```

```

QY      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGly 160
DB      1003 AGCGGAAGTAGAGACAGCTCTGGCTTAAATTTTGGCCGGATGGGGGATTATCGAAATGAA 1062
QY      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB      1063 ACCTTGACAGCAACACCGCGCGACAGCTCCCTTTCTTCCCACTTGTTACAGACGTAATTT 1122
QY      181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
DB      1123 TTCTGCGCGGCATAGACGTTGTTCTTCCTGCGCAATGCCGATACAGATGTTGTTATTAC 1182
QY      201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHisLeuTyrAsnAlaGluThr 220
DB      1183 ATCGACGATTCGAGACGATTCGCAACGAAACCGAATGCACTATCAATGCCGAACA 1242
QY      221 LeuLysAlaGlnThrLysLeuGlyuThrPheAlaValAspArgThrAsnLysLeuLeu 240
DB      1243 CTGAAGCCCAAAACAAACTGGAATATTTCCAGCTAGCAGAACCAATAAAAATTTGCTC 1302
QY      241 IleLysProLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTrpMet 260
DB      1303 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTTAAAGAAATACGCAATGTGATG 1362
QY      261 GlyProTyrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer 280
DB      1363 GGGCGGTATAAAGTATAGCAAAAGGATCAACCGACGGAAGATTATGTGATTTCTCC 1422
QY      281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300
DB      1423 GATATCCGGCCATACGCGCAATCATACGCGTAACTCCGCCCATCCGTAGAGCGTGAATAC 1482
QY      301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320
DB      1483 AGTCATAGGGGATGATGATACGCGATGACGACAGTGCACACATGACAAAGGCCAAGCT 1542

RESULT 2
ABZ37917
ID      ABZ37917 standard; DNA; 960 BP.
XX      AC      ABZ37917;
XX      DT      07-MAR-2003 (first entry)
XX      DE      N. gonorrhoeae nucleotide sequence SEQ ID 423.
XX      KM      Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX      OS      Neisseria gonorrhoeae.
XX      PN      WO200279243-A2.
XX      PD      10-OCT-2002.
XX      PE      12-FEB-2002; 2002WO-IB002069.
XX      PR      12-FEB-2001; 2001GB-00003424.
XX      PA      (CHIR-) CHIRON SPA.
XX      PI      Fontana MR, Pizza M, Maignani V, Monaci E;
XX      DR      WPI; 2003-058415/05.
XX      DR      P-PSDB; ABP76947.
XX      FT      New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX      PS      medicament for treating or preventing N. gonorrhoeae infection.
XX      PS      Disclosure; Page 220; 815pp; English.
XX      CC      The present invention relates to proteins from Neisseria gonorrhoeae.
XX      CC      Also disclosed are the nucleic acid molecules encoding the proteins and
XX      CC      antibodies that specifically bind to the proteins. The composition

```



CC completing the protein, nucleic acid or antibody is useful for the  
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
 CC infection, this may be in the form of a vaccine or gene therapy.  
 CC Sequences given in records AB237706-AB242016 represent nucleic acid  
 CC molecules of the invention

XX Sequence 960 BP; 275 A; 241 C; 225 G; 219 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	2,57e-158	Length:	960
Score:	1641.00	Matches:	319
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	99.76%	Indels:	0
DB:	10	Gaps:	0

US-10-617-835-4 (1-320) x AB237917 (1-960)

```

QY 1 MetATGAlaArgLeuLeuIleProIleuPheSerValPheIleuSerAlaCysGly 20
DB 1 ATGCGGACGGCTGCTGATACCTATTCTTTTTCAGTTTATTATTCCTGCGGG 60
QY 21 ThrLeuThrGlyIleProSerHisGlyGlyIleYsArgPheAlaValGluGluLeu 40
DB 61 ACATGACAGCATATTCATCGCATGCGGAGCGCAACCGCTTCGCGTCAACAAGACTT 120
QY 41 ValAlaIleSerAlaArgAlaAlaValYsAspMetAspLeuGlnAlaLeuHisGlyArg 60
DB 121 GTGGCGGCTTTCGACAGAGCTCCGTTAAAGACATGATTTACAGGCATTACACGACGA 180
QY 61 LysValAlaLeuTyrlleAlaThrMetGlyYsAspGlnGlySerIleuThrGlyGly 80
DB 181 AAAGTTGCACTTGTACATTGCACTAATGGCGACAAAGTTTCAGGCAAGTTGACAGGGGT 240
QY 81 ArgTYrSerIleAspAlaLeuIleArgGlyGlyTYrlleAsnSerProAlaValArgThr 100
DB 241 CGCTACTCATTTGATGACACTGATTCGGCGGCGAATCATAAACAGCCCTGCGGACCC 300
QY 101 AspTYrThrTYrProArgTYrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
DB 301 GATTACACCTATCCCGTTACGAAACCAACCGCTGAAACAAACATCAGCGGTTGACGGGT 360
QY 121 LeuThrThrSerLeuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB 361 TTAAACCACTTCTTAACTACACTTAATGCCCTGCACTCTCCGCAACCAATCAAGCGGT 420
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTYrArgAsnGlu 160
DB 421 AGCGGAAGAGAGCAGCTGGGCTTAAATATTGGCGGATGGGGGATTATCGAAATGA 480
QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB 481 ACCTTGACGACCAACACCGCGGACACCTGCTTCTTCCACCTGTCGACAGCCGATATT 540
QY 181 PheLeuArgGlyIleAspValYsAspProAlaAsnAlaAspThrAspValPheIleAsn 200
DB 541 TTCCTGCGGCGCATGACGTGTCTCTCTGCAATGCCGATACAGATGTGTTATTAC 600
QY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTYrAsnAlaGluThr 220
DB 601 ATCGACGATATTGGAACGATACGCAACGAAACCGAAATCAGCTATACATCCGAAACA 660
QY 221 LeuValAlaGlnThrIleLeuGluTYrPheAlaValAspArgThrAsnValYsLeuLeu 240
DB 661 CTGAAGACCCCAACAAACCTGCAATATTTCGACAGTACAGAACCAATAAAAATTCGCTC 720
QY 241 IleLeuProLeuThrAsnAlaPheGluAlaIleTYrYsGluAsnTYrAlaLeuTYrPhe 260
DB 721 ATCAAAACCAAAACCAATCGCTTTGAGCTGCTTAAAGAAATTAACCATTCGATG 780
QY 261 GlyProTYrYsValSerIleGlyIleYsProThrGluGlyLeuMetValAspPheSer 280
DB 781 GGGCGGTATTAAGTAAAGCAAGATCAACCGACGAGAGATGATGATTCATTTCTCC 840
  
```

```

QY 261 AspIleArgProTYrGlyYsAsnHisThrGlyYsAsnSerAlaProSerValGluAlaAspAsn 300
DB 841 GATATCCACCAACATACGGAATCATACGGGTAACTCCGCCCATTCGTAGAGCTCATTAAC 900
QY 301 SerHisGlyGlyTYrGlyTYrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320
DB 901 ACTCATGAGGGGTATGATACGATACGATGACAGTGCACAACTACAGAGGCAACT 960
  
```

#### RESULT 3

AAZ12008  
 ID AAZ12008 standard; DNA; 963 BP.

AC AAZ12008;

DT 08-OCT-1999 (first entry)

DE Neisseria gonorrhoeae complete ORF15 sequence.

KM Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
 KM treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.

OS Neisseria gonorrhoeae.

PN W0924578-A2.

PD 20-MAY-1999.

PF 09-OCT-1998; 98WO-IB001665.

PR 06-NOV-1997; 97GB-00023516.

PR 14-NOV-1997; 97GB-00024190.

PR 18-NOV-1997; 97GB-00024386.

PR 27-NOV-1997; 97GB-00025158.

PR 10-DEC-1997; 97GB-00026147.

PR 14-JAN-1998; 98GB-00000759.

PR 01-SEP-1998; 98GB-00019016.

XX (CHIR-) CHIRON SPA.

PA Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;

PI WPI: 1999-327407/27.

DR P-PSDB; AA38540.

PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for

PS diagnosis, treatment and prevention of infection.

XX Claim 9; Page 100; 524pp; English.

CC Nucleotide sequences AAZ11972-212358 represent open reading frames (ORFs)

CC of Neisseria meningitidis and N. gonorrhoeae which encode antigenic

CC proteins (see AA38499-Y38944). The antigenic proteins, their fragments,

CC their nucleic acids and antibodies are used for diagnosis, prevention (as

CC vaccines) or treatment of Neisseria infections, such as meningitis,

CC septicemia and gonorrhea. Both organisms are closely related. Fragments

CC of the nucleic acids are useful as hybridisation probes and antisense

CC reagents

SQ Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	2,58e-158	Length:	963
Score:	1641.00	Matches:	319
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	99.76%	Indels:	0
DB:	2	Gaps:	0

US-10-617-835-4 (1-320) x AAZ12008 (1-963)

```

QY 1 MetATGAlaArgLeuLeuIleProIleuPheSerValPheIleuSerAlaCysGly 20
  
```

```

Db      1 ATGCGGACGCGCTGTGATACCTATCTTTTTCAGTTTATTTTATCCGCTCGCGG 60
QY      21 ThrLeuThrGlyIleProSerHisGlyIleValysArgPheAlaValGluGlnIleu 40
Db      61 ACACAGACAGGATTCATCGCATGCGGAGCAAAACGCTTCGCGTGAACAACAACTT 120
QY      41 ValAlaIleSerAlaArgAlaAlaValysAspMetAspLeuGlnAlaLeuHisGlyArg 60
Db      121 GTGGCGGCTTCTGCAGAGCTGCGTTAAAGCAATGAGATTACAGCATTTACACGACGA 180
QY      61 LysValAlaLeuTyrIleAlaThreGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Db      181 AAAGTTGATTTGATTCATTCACATATGCGGCGACCAAGGTTTCAGGCAAGTTTGACAGGGGGT 240
QY      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
Db      241 CGCTACTCATTTGATGACCTGATTCGCGGCAATTCATTAACAGCCCTGCGTCCGACCC 300
QY      101 AspTyrThrTyrProArgTyrGlyIleuThrAlaGluThrThrSerGlyIleuThrGly 120
Db      301 GATTACACCTATCCGCTTACGAACCAACCGCTGAACAACATCAAGCGGTTTGACGCGGT 360
QY      121 LeuThrThrSerLeuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db      361 TTAACCACTTCTTATCTACACTTATATCCCTGCACCTCGCGACCAATCAAGACCGT 420
QY      141 SerGlySerArgSerSerLeuGlyIleuAsnIleGlyIleMetGlyAspTyrArgAsnGlu 160
Db      421 AGCGGAAGTACGACGATCTGGGCTTAAATTTGCGGGAATGGGGGATTTACGAATGAA 480
QY      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisIleValGlnThrValPhe 180
Db      481 ACCTTGACGACCAACCGCGCGACACCTGCTTTTCCACTTGTCGACAGCGTATTT 540
QY      181 PheLeuArgGlyIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541 TTCCTCGCGGCAATAGACGTGTCTCTCCCAATGCCATACGATGTGTTATTTAAC 600
QY      201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisIleuTyrAsnAlaGluThr 220
Db      601 ATCAACGATATTCGAGACGATACCGCAACGAAACCGAAATGCACTTATACATGCCGAACA 660
QY      221 LeuLysAlaGlnThrIleuGlyIleuGlyPheAlaValAspArgThrAsnIleuLeu 240
Db      661 CTGAAGGCCCAAACTGGAATATTTCCAGATGACAGAACCAATTAATAAATAAATTTGCTC 720
QY      241 IleLysProLysThrAsnAlaPheGluAlaIleTyrIleGlyIleuAsnTyrAlaLeuTrpMet 260
Db      721 ATCAAAACCAAAACCAATGCGTTGAAAGCTCCCTATTAAGAAAATTTACGCAATTGTGATG 780
QY      261 GlyProTyrLysValSerIleGlyIleLysProThrGluGlyIleuMetValAspPheSer 280
Db      781 GGGCCGATTAAGTAAGCAAGAAAGATCAACCGACGGAAGGATTTGATGATTTCTCC 840
QY      281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300
Db      841 GATATCCAAACCATACCGGCAATCATACGGGTAACTCCGCCCATCGTGAAGCGTGATTAAC 900
QY      301 SerIleGlyGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyIlePro 320
Db      901 AGTCATGAGGGGTATGATACAGCGATGAACAGTGCACAACTAGACAGAGGGCAACCT 960

```

## RESULT 4

AAZ54614  
AAZ54614 standard; DNA, 963 BP.

AC AAZ54614;

DT 15-SEP-2003 (revised)  
DT 21-MAR-2000 (first entry)

DE Neisseria gonorrhoeae ORF 406 partial DNA sequence SEQ ID NO:3111.

XX

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy; ds.

OS Neisseria gonorrhoeae.

PN W0957280-A2.

PD 11-NOV-1999.

XX 30-APR-1999; 99WO-US009346.

PR 01-MAY-1998; 98US-0083758P.

PR 31-JUL-1998; 98US-0094869P.

PR 02-SEP-1998; 98US-0098994P.

PR 02-SEP-1998; 98US-0099062P.

PR 09-OCT-1998; 98US-0103749P.

PR 09-OCT-1998; 98US-0103794P.

PR 09-OCT-1998; 98US-0103796P.

PR 25-FEB-1999; 99US-0121528P.

XX (CHIR ) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

P1 Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M,  
Pi Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M,  
Pi Tettelin H, Venter JC,

DR WPI: 2000-062150/05.

XX P-PSDB; AAY75812.

PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.

XX Example 1; Page 109; 1453bp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
XX represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
XX and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
XX PCR primers used in the exemplification of the present invention. The  
XX polypeptides, the polynucleotides, antibodies and compositions of the  
XX invention can be used as vaccines, as diagnostic reagents, and as  
XX immunogenic compositions. The polypeptides can be used in the manufacture  
XX of medicaments for treating or preventing infection due to Neisserial  
XX bacteria (e.g. meningitis and septicemia), to detect the presence of  
XX Neisseria bacteria, or to raise antibodies. They may also be used to  
XX screen for agonists or antagonists, which may themselves have use as  
XX antibacterial agents. The polynucleotides of the invention may also be  
XX used in gene therapy protocols. (Updated on 15-SEP-2003 to standardise OS  
XX field)

XX SQ Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;  
XX

## Alignment Scores:

Pred. No.:	2,58e-158	Length:	963
Score:	1641.00	Matches:	319
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	99.76%	Indels:	0
DB:	3	Gaps:	0

US-10-617-835-4 (1-320) x AAZ54614 (1-963)

QY 1 MetArgAlaArgLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20

Db 1 ATGCGGACAGGCTGTGATACCTATCTTTTTCAGTTTATTTTATTCCTCGCGGG 60

QY 21 ThrLeuThrGlyIleProSerHisGlyIleValysArgPheAlaValGluGlnIleu 40

Db 61 ACACAGACAGGATTCATCGCATGCGGAGCAAAACGCTTCGCGTGAACAACAACTT 120

QY 41 ValAlaIleSerAlaArgAlaAlaValysAspMetAspLeuGlnAlaLeuHisGlyArg 60

```

Db      121 GTGGCGGCTTCTCCAGAGCTGCCGTTAAAGACATGATTTACAGGCATTACAGGAGCA 180
Qy      61 LysValAlaLeuTyrTlIleAlaThMeCglYAspGlnGlySerGlySerLeuThrglyGly 80
Db      181 AAAGTTGATTTGATTCATTCGCACTATGGCGGACCAAGGTTGAGGAGATTGACAGGGGGT 240
Qy      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrTlIleAspSerProAlaIleArgThr 100
Db      241 CGCTACCTCATTTGATGACTGATTCGGCGGAAATCATTAACAGCCCTGCGCTCGCACCC 300
Qy      101 AspTyrThrTyrPrcArgGlyTyrTlIleThraIaGluThrThSerGlyGlyLeuThrgly 120
Db      301 GATTACACTATCCGCGTTACGAACCCGCTGAACCAACATCAGCGCGTTTGAACGGGT 360
Qy      121 LeuThrThrSerLeuSerThrIleuAsnAlaProAlaLeuSerArgTlIleSerAspGly 140
Db      361 TTAAACACTCTTTTATCTACTACTATATGCCCCCTGCACCTCCGACCAACCAATCAGACGGT 420
Qy      141 SerGlySerArgSerSerIleuGlyLeuAsnIleGlyGlyMetGlyAspTyrTrgAsnGlu 160
Db      421 AGCGGAGAGAGAGAGAGCTGGGCTTAAATATGGCGGATGGGGATTATCGAAATGAA 480
Qy      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHleuValGlnThrValPhe 180
Db      481 ACCTTGACGAGCAACCCCGCGCACACTGCTTTCTTCCCACTTGTTGACAGCCGTAATT 540
Qy      181 PheLeuArgGlyTlIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541 TTCCTGCGCGCATAGACCTGTTCTCTCGCCAATGCGCATACAGATGTGTTATTAAAC 600
Qy      201 IleAspValIlePheGlyThrIleArgAsnArgTlIleGlyMetHleuIleuThrAsnIleGluThr 220
Db      601 ATCGACGATTTGCGAACGATACGCAACCAACCGAAATGACACTTATACATGCCAACA 660
Qy      221 LeuIleValIleGlnThrIleuGlyTyrPheAlaValIleAspArgThrAsnIleIleuLeu 240
Db      661 CTGAAGGCCCAACAACTGGAATTTTCCGAGTATGACAGAACCAATTAATAATTGCTC 720
Qy      241 IleIleProlYsThrAsnAlaPheGluAlaIleTlIleGlyAsnTlIleAlaLeuTlIlePhe 260
Db      721 ATCAAAACCAAAACCAATGCGTTGAGCTGCTTAAAGAAATATACCATTTGCGATG 780
Qy      261 GlyProlYsIleAspValIleSerIleGlyTlIleIleProlYsIleuMetValIlePheSer 280
Db      781 GGGCCGATTAAGTAAAGAACCAAGAAATCAAAACCGACGAAAGTATGTCGATTTCTCC 840
Qy      281 AspIleArgProlYsIleAsnIleThrgIleAsnSerAlaProlSerValGluIleAspAsn 300
Db      841 GATATCCAAACATACGCGCATATACGGGTAACTCCGCCCATCCGTAAGGCTGATAAC 900
Qy      301 SerHleGluGlyTyrGlyTyrIleSerAspGluAlaValIleArgIleHleArgGlnGlyGlnPro 320
Db      901 AGTCATGAGGGGTATGATGATACGATGACGATGACGACGACGACCAATGACCAAGGGCAACCT 960

```

RESULT 5  
AAZ53688  
ID AAZ53688 standard; DNA; 963 BP.

AAZ53688;  
AC  
XX 15-SEP-2003 (revised)  
DT 21-MAR-2000 (first entry)  
XX  
DE Neisseria gonorrhoeae ORF 406 partial DNA sequence SEQ ID NO:1325.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy; ds.  
XX  
OS Neisseria gonorrhoeae.  
XX  
PN W09957280-A2.  
XX

```

PD      11-NOV-1999.
XX
PF      30-APR-1999; 99WO-US009346.
XX
PR      01-MAY-1998; 98US-0083758P.
XX      31-JUL-1998; 98US-0094869P.
PR      02-SEP-1998; 98US-0098994P.
PR      02-SEP-1998; 98US-0099062P.
PR      09-OCT-1998; 98US-0103749P.
PR      09-OCT-1998; 98US-0103794P.
PR      09-OCT-1998; 98US-0103796P.
PR      25-FEB-1999; 99US-0121528P.
XX
PA      (CHIR) CHIRON CORP.
PA      (GENO-) INST GENOMIC RES.
XX
PI      Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
PI      Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI      Tettelin H, Venter JC;
XX
DR      WPI, 2000-062150/05.
DR      P-PsDB; AAY74926.
XX
PT      Novel Neisserial polypeptides predicted to be useful antigens for
PT      vaccines and diagnostics.
XX
PS      Claim '7, Page 728; 1453pp; English.
XX
CC      AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC      represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
CC      and polypeptides; AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC      PCR primers used in the exemplification of the present invention. The
CC      polypeptides, the polynucleotides, antibodies and compositions of the
CC      invention can be used as vaccines, as diagnostic reagents, and as
CC      immunogenic compositions. The polypeptides can be used in the manufacture
CC      of medicaments for treating or preventing infection due to Neisserial
CC      bacteria (e.g. meningitis and septicemia), to detect the presence of
CC      Neisseria bacteria, or to raise antibodies. They may also be used to
CC      screen for agonists or antagonists, which may themselves have use as
CC      antibacterial agents. The polynucleotides of the invention may also be
CC      used in gene therapy protocols. (Updated on 15-SEP-2003 to standardise OS
CC      field)
XX
SQ      Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;

```

## Alignment Scores:

Pred. No.: 2,598-158 Length: 963  
Score: 1641.00 Matches: 319  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.69% Mismatches: 0  
Query Match: 99.76% Indels: 0  
DB: 3 Gaps: 0

US-10-617-835-4 (1-320) x AAZ53688 (1-963)

```

Qy      1 MetArgAlaArgLeuIleIleProIleuPheSerValPheIleLeuSerAlaCysGly 20
Db      1 ATGGCGGACGCGCTGCTGATACCTATCTTTTTCAGTTTATTTATTCGCCCTGCGG 60
Qy      21 ThrLeuThrGlyIleProlSerHleGlyGlyGlyIleAspPheAlaValIleGlnGlu 40
Db      61 ACACGACAGAGTATTCATTCGCAATGCGGACGAAACCTTCGCGCTGCAACAGAACTT 120
Qy      41 ValAlaIleSerAlaArgAlaAlaValIleAspMetAspLeuGlnAlaLeuHleGlyArg 60
Db      121 GTGGCGGCTTCTGCGACAGCTGCGCTTAAACATGATGATTTACAGGCAATTACACCGAGCA 180
Qy      61 LysValAlaLeuTyrTlIleAlaThMeCglYAspGlnGlySerGlySerLeuThrglyGly 80
Db      181 AAAGTTGATTTGATTCATTCGCACTATGGCGCAACAGGTTTCAGGCGATTGACAGGGGGT 240
Qy      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrTlIleAsnSerProAlaValArgThr 100

```

```

Db      241 CGTACTCATTCATGATGACGATTCGGCGGCAATACATAAACAGCCCTCGCGCAC 300
Qy      101 AspTYrThrTYrProAlaTYrGluThrThraAlaGluThrThrSerGlyGlyLeuThrGly 120
Db      301 GATTACACCTTATCCGCGTTACGAAACCAACCCCTGTAACAAATCAGCGGGGTTTGAACGGGT 360
Qy      121 LeuThrThrSerLeuSerThrThraAlaAlaProAlaLeuSerArgThrGlySerAspGly 140
Db      361 TTAACCACTTCTTATCTACACTTATACCTCCCTGACCTCGCGACCAACCAATCAGACGGT 420
Qy      141 SerGlySerArgSerSerLeuGlyLeuAlaSerGlyGlyMetGlyAspTYrArgAsnGlu 160
Db      421 AGCGGAAGTAGAGAGAGCTGGGCTTAATATGTCGGGATGGGGGATTAATGCAATGAA 480
Qy      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerThrAlaGluThrValPhe 180
Db      481 ACCTTGAGACCAACCCGCGGACACTGCTCTTCTTCCACTGGTGCAAGCCGATATT 540
Qy      181 PheLeuArgGlyIleAspValValSerProAlaAlaSerThrAspValPheIleAsn 200
Db      541 TTCCTCGCGCGCATGACGTTGTTCTCTCGCAATGCGATACAGATGCTTTATTATAC 600
Qy      201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTYrAsnAlaGluThr 220
Db      601 ATCGACGATTCGGAAACGATACGCAACCAACCGAAATGCACTATACAAATGCCGAAACA 660
Qy      221 LeuIysAlaGluThrIleValLeuGluThrPheAlaValAspArgThrAsnIleValLeu 240
Db      661 CTGAAAGCCCAACCAAACTGAAATATTCGACAGACAGAAACCAATTAATAATTTGCTC 720
Qy      241 IleIysProIysThrAsnAlaPheGluAlaIleAlaTYrIleValAsnTYrAlaLeuTrpMet 260
Db      721 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTCTATAAAGAAATTAACGCAATTGGAGATG 780
Qy      261 GlyProTYrIleValSerIleGlyIleIysProThrGluGlyLeuMetValAspPheSer 280
Db      781 GGGCGCATTAAGTAAGCAAAAGAAATCAACCGAAGGATTTGATGCTGATTTCTCC 840
Qy      281 AspIleArgProTYrGlyAsnHisThrGlyValSerSerAlaProSerValGluAlaAspAsn 300
Db      841 GATATCCAAACATACGCGAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATAC 900
Qy      301 SerHisGluGlyTYrGlyTYrSerAspGluAlaValAlaArgIleHisArgGluGlyLeuPro 320
Db      901 AGTCATGAGGGGTATGATACAGCATGACGATGACAGCAACATAGCAAGGGCAACCT 960
RESULT 6
AAAB1297
ID      AAAB1297 standard; DNA; 963 BP.
XX
XX      AAA81297;
XX
XX      15-SEP-2003 (revised)
DT      04-DEC-2000 (first entry)
XX
DE      N. gonorrhoeae partial DNA sequence g406.seq SEQ ID NO:1036.
XX
XX      Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW      antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW      Meningococcus B; MenB; ds.
XX
XX      Neisseria gonorrhoeae.
OS
XX      MO200022430-A2.
PN
XX      20-APR-2000.
PD
XX      08-OCT-1999; 99WO-US023573.
PF
XX      09-OCT-1998; 98US-0103794P.
PR      30-APR-1999; 99US-0132068P.
XX
XX      (CHIR ) CHIRON CORP.
PA

```

```

XX      Frazer CM, Hickey B, Peterson J, Tertelin H, Venter JC;
PI      Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI      Rappuoli R, Pizza M;
XX      WPI; 2000-318079/27.
DR      P-PSDB; AAB25657.
XX
XX      Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT      in the diagnosis and treatment of N. meningitidis infection and other
PT      Neisserial infections, for example, N.gonorrhoea.
PS      Claim 9; Page 108; 1760pp; English.
XX
XX      The present invention describes methods of obtaining immunogenic proteins
CC      from Neisseria genomic sequences. AAAB1453 to AAAB2414 represent
CC      specifically claimed Neisseria meningitidis genomic DNA sequences;
CC      AAAB1260 to AAAB1303 and AAB25620 to AAB25663 represent Neisseria DNA
CC      sequences and their corresponding proteins; AAAB1254 to AAAB1259 and
CC      AAAB1304 to AAAB1321 represent PCR primers used in the isolation of
CC      Neisseria meningitidis DNA sequences; and AAAB1322 to AAAB1452 represent
CC      Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC      used in the exemplification of the present invention. The nucleic acid
CC      sequences, protein sequences, and antibodies against them, can be used in
CC      the manufacture of a composition. The composition can be used as a
CC      medicament (or in the manufacture of a medicament) for treating,
CC      preventing or diagnosing infection due to Neisserial bacteria. For
CC      example, some of the identified proteins could be components of vaccines
CC      against Meningococcus B; against all serotypes; and/or against all
CC      pathogenic Neisseriae. Identification of sequences from the bacterium
CC      will also facilitate production of biological probes, particularly
CC      organism-specific probes. Attempts to make efficacious Meningococcus B
CC      vaccines have also been tried but none have successfully overcome
CC      antigenic variability. The provision of further, complete sequences may
CC      provide an opportunity to identify secreted or surface exposed proteins
CC      that may be presumed targets for the immune system and which are not
CC      antigenically variable or at least more conserved than other more
CC      variable regions. (Updated on 15-SEP-2003 to standardise OS field)
XX
XX      Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 2,58e-158 Length: 963
Score: 1641.00 Matches: 319
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.69% Mismatches: 0
Query Match: 99.76% Indels: 0
DB: Gaps: 0
US-10-617-835-4 (1-320) x AAAB1297 (1-963)
Qy      1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db      1 ATGCGGGACAGCGCTGCGATACCTATCTTTTTCAGTTTATTATTATCCGCTCGGG 60
Qy      21 ThrLeuThrGlyIleProSerHisGlyGlyIleValArgPheAlaValGluGluLeu 40
Db      61 ACACTGACAGGTATTCATCGCATGCGGAGGCAAAAGCTTCGGGTCGAACAAAGAACTT 120
Qy      41 ValAlaAlaSerAlaArgAlaAlaValIleAspMetAspLeuGluAlaLeuHisGlyArg 60
Db      121 GTGGCCGCTTCTGCGCAAGCTGCGCTTAAGACATGATTTACAGCATTTACAGGACGA 180
Qy      61 LysValAlaLeuTYrIleAlaThrMetGlyAspGluGlySerGlySerLeuThrGlyIle 80
Db      181 AAAGTTCATTTGATGACGATGACCTATGAGCGACCAAGCTTCAGGCACTTTGACAGGGGCT 240
Qy      81 ArgTYrSerIleAspAlaLeuIleArgGlyGlyTYrIleAsnSerProAlaValArgThr 100
Db      241 CGTACTCATTCATGATGACGATTCGGCGGAAATACATTAACAAACCCCTCGCGCAC 300
Qy      101 AspTYrThrTYrProAlaTYrGluThrThraAlaGluThrThrSerGlyGlyLeuThrGly 120

```

```

Db      301 GATTACACCTATCCGGGTTACGAAACACCGCTGAAACAAACATCAGCCGGTTGACGGGT 360
Qy      121 LeuThrThrSerLeuSerThrLeuAana1aPro1aLeuSerArgThrGlnSerAapGly 140
Db      361 TTACACACTTCTTATCTACCTAATATGCCCCGCACTCCGGGCAACCCCAATCAGACGGGT 420
Qy      141 SerGlySerArgSerSerLeuGlyLeuAan11GlyGlyMetGlyAapThrArgAanGlu 160
Db      421 AGCGGAAGAGAGGAGCGCTGGCTTAAATATGCGGGATGGGAGATTATCCAAATGAA 480
Qy      161 ThrLeuThrThrAsnProArgAapThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      481 ACCTTACACACCAACCCGCGGCACTGCTTCTTCCCACTTGGTGCAGACCGATTAT 540
Qy      181 PheLeuArgGly11LeaApyVala1SerPro1aAana1aAapThrApyValPhe1LeaA 200
Db      541 TTCCTGCGGGCATACGCTGTTTCTCCTGCAATGCCGATACAGATGTTATTATAC 600
Qy      201 11LeaApyValPheGlyThr11LeaAana1aGThrGluMetHisLeuThrAan1aGluThr 220
Db      601 ATCGACGATTCGGAAACGATACGCAACAGAACCGAAATGCACCTATACATGCGCAAAAC 660
Qy      221 LeuLysAlaG11ThrLysLeuGlu1uTyrrPheAlaVala1AapArgThrAsnLysLysLeu 240
Db      661 CTGAAAGCCCAACAACTGGAATTTTCGCGATGACAGAACCAATTAATAATTGCTC 720
Qy      241 11LeuApyProLysThrAana1aPheGlu1a1a1aTyrrLysGluAanTyrrAlaLeuThrMet 260
Db      721 ATCAAAACCCAAACCAATCGTTGAAAGCTGCTTAAAGAAATTAACCATTTGTGGAG 780
Qy      261 GlyProTyrrLysValSerLysGly11LeuApyProThrGlnGlyLeuMetVala1AapPheSer 280
Db      781 GGGCGGTATTAAGTAAGCAAGCAACCAACCGAAGGATGATGATGATTCCTCC 840
Qy      281 Aap1LeaArgProTyrrLysAanHisThrGlyAanSer1aProSerValGlu1aAaPasn 300
Db      841 GATATCCAAACCATACGGCAATCATACGGGTAACTCCGCCCATCCCTAAGGCTGATAC 900
Qy      301 SerHisGluGly1TyrrGlyTyrrSerAapGlu1aVala1ArgGlnHisArgGlnGlyInPro 320
Db      901 AGTCATGAGGGGATGATGATACAGCGATGAAAGCGTGCGAACATATGACAAAGGCAACCT 960

```

## RESULT 7

AAF21582  
ID AAF21582 standard; DNA; 963 BP.

```

XX      XX      AAF21582;
AC      AC      15-SEP-2003 (revised)
XX      XX      DT      13-MAR-2001 (first entry)
XX      XX      DE      N. gonorrhoeae partial DNA sequence g406.seq SEQ ID NO.76.
XX      XX      KM      Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX      XX      KM      diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX      XX      OS      Neisseria gonorrhoeae.
XX      XX      PN      WO200066791-A1.
XX      XX      PD      09-NOV-2000.
XX      XX      PF      08-MAR-2000; 2000WO-US005928.
XX      XX      PR      30-APR-1999; 99US-0132068P.
XX      XX      PR      08-OCT-1999; 99WO-US023573.
XX      XX      PR      28-FEB-2000; 2000GB-00004695.
XX      XX      (CHIR ) CHIRON CORP.
XX      XX      PA      (GENO-) INST GENOMIC RES.
XX      XX

```

```

PI      Pizza M, Hickey E, Peterson J, Tetrelin H, Venter JC;
PI      Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI      Rapinoli R, Frazer CM, Grandi G;
XX      DR      MPI; 2000-647603/62.
XX      DR      P-PSDB; AAB58587.
XX      PT      Neisseria meningitidis B full length genome sequence and open reading
XX      PT      frames are used to detect, treat and prevent Neisserial infections.
XX      PS      Example 1; Page 109; 692pp; English.
XX      CC      The present invention describes the full length genome of Neisseria
XX      CC      meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
XX      CC      represent fragments of the NMB genomic sequence, as the sequence was too
XX      CC      long to go in a record on its own it was split into 8 sequences which
XX      CC      overlap each other at the beginning and end of each sequence by 49980 bp
XX      CC      (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
XX      CC      AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
XX      CC      AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
XX      CC      given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
XX      CC      primers which are used in the exemplification of the present invention.
XX      CC      The NMB genome and fragments from it have antibacterial activity, and can
XX      CC      be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
XX      CC      and/or antibodies which binds to the proteins can be used in compositions
XX      CC      for treating or preventing infection due to Neisserial bacteria or as a
XX      CC      diagnostic reagent for detecting the presence of Neisserial bacteria or
XX      CC      of antibodies raised to Neisserial bacteria. Computers, computer memory,
XX      CC      computer storage medium or computer databases can be used in a search to
XX      CC      identify open reading frames (ORFs) or coding sequences within the NMB
XX      CC      genome. The DNA sequences provide further opportunities to find antigenic
XX      CC      or immunogenic proteins which are more effective in vaccines than the
XX      CC      outer membrane proteins currently used. (Updated on 15-SEP-2003 to
XX      CC      standardise OS field)
XX      SQ      Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Score: 2,586-158 Length: 963
XX      Percent Similarity: 1641.00 Matches: 319
XX      Best Local Similarity: 100.00% Conservative: 1
XX      Query Match: 99.76% Indels: 0
XX      DB: 3 Gaps: 0
XX
XX      US-10-617-835-4 (1-320) x AAF21582 (1-963)
Qy      1 MetArgAlaArgLeuLeu11ePro11eLeuPheSerValPhe1eLeuSer1aCysGly 20
Db      1 ATGCGGGACGCGCTGCTGATTAATCTTTTTCAGTTTATTTATTCCTCCGCGGG 60
Qy      21 ThrLeuThrGly11eProSerHisGlyGlyLysArgPheAlaVala1GlnGlnLeu 40
Db      61 ACATGACAGGATTCATTCGATCGCGAGGCAACCGCTTCGCGTGCMAACAACACTT 120
Qy      41 ValAlaAlaSerAlaArgAlaAlaVala1LysAapMetAapLeuGlnAlaLeuHisGlyArg 60
Db      121 GTGGCCGCTTCGCGAGAGCTGCGTTAAAGCAATGATTTACAGCATTTACACGACCA 180
Qy      61 LysValAlaLeuTyrr1eAla1aThrMetGlyAapGlnGlySerGlySerLeuThrGly 80
Db      181 AAAGTTGATTTGATTCATTCGCGAGGCAACAGGTTGACGAGTTGACAGGGGT 240
Qy      81 ArgTyrrSer11eAapAlaLeu11eArgGlyGlyTyrr11eAanSerPro1aVala1ArgThr 100
Db      241 CGCTATCTCATTTGATGACATGATTCGCGCGAATACATTAACAGCCCTGCCCGCACCC 300
Qy      101 AapTyrrThrTyrrProArgTyrrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
Db      301 GATTACACCTATCCGGGTTACGAAACACCGCTGAAACAAATCAGCGCGTTGACGGGT 360
Qy      121 LeuThrThrSerLeuSerThrLeuAana1aPro1aLeuSerArgThrGlnSerAapGly 140

```

```

Db      361  TTAACCACTTTTATCTACTACTTAATGCCCCCTGCAGCTCTCGCGACCAACCAACGAGT 420
Qy      141  SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspPheArgAsnGlu 160
Db      421  AGCGGAAGAGAGAGAGCTGGGCTTAATAATTTGGGGATGGGGATATTCGAATGAA 480
Qy      161  ThrLeuThrThrAsnProArgSerPheThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      481  ACCTTGACGACCAACCCCGCGACACTGGCTTTCTTTCCCACTTGTGTGACGCGTATTT 540
Qy      181  PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541  TTCCGCGCGGCGATAGACGTTGTTCTCTCTGCCAATGCCGATACAGATGTTATTTAAC 600
Qy      201  IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyAsnAlaGluThr 220
Db      601  ATCGACGATTTGGAAACGATACGCAACGAAACCGAAATGCACTTATACATGCCAACA 660
Qy      221  LeuIysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysIleLeu 240
Db      661  CTGAAGCCCAAAACAACCTGGATATTTCCGAGTAGACAGAACCAATTAATAATTTGCTC 720
Qy      241  IleLysProLysThrAsnAlaPheGluAlaAlaIleTyrLysGluAsnTyrAlaLeuTyrMet 260
Db      721  ATCAAAACCCAAACCAATGCGTTTGAGCTGCTTAAAGAAATTAACGCAATGTGGATG 780
Qy      261  GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSer 280
Db      781  GGGCGGTATTAAGTAAAGCAAGAGAAATCAAAACCGAAGAGATTAATGATTCATTTCTCC 840
Qy      281  AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300
Db      841  GATATCCAAACCATACGGCATATCATCGGGTAACTCCGCCCATTCGTAGAGGCTGATTAAC 900
Qy      301  SerHisGluGlyTyrGlyTyrSerAspGluAlaValAlaArgGlnHisArgGlnGlyGlnPro 320
Db      901  AGTCATGAGGGGTATGATGATACAGCGATGAAAGCATGACAAACATTAAGCAAGGGCAACT 960

RESULT 8
AAZ12006
ID      AAZ12006 standard; DNA; 963 BP.
AC      AAZ12006;
CT      08-OCT-1999 (first entry)
DE      Neisseria meningitidis strain B complete ORF15 sequence.
KW      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
        treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.
XX      OS
        Neisseria meningitidis.
XX      EN
        W09924578-A2.
XX      PD
        20-MAY-1999.
XX      PF
        09-OCT-1998; 98WO-IB001665.
XX      PR
        06-NOV-1997; 97GB-00023516.
        14-NOV-1997; 97GB-00024190.
        18-NOV-1997; 97GB-00024386.
        27-NOV-1997; 97GB-00025158.
        10-DEC-1997; 97GB-00026147.
        14-JAN-1998; 98GB-00000759.
        01-SEP-1998; 98GB-00019016.
XX      PA
        (CHIR-) CHIRON SPA.
XX      PI
        Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX      WP1; 1999-327407/27.
XX      DR
        P-PSDB; AAY38538.

```

```

XX      XX
PT      Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT      diagnosis, treatment and prevention of infection.
PS      Claim 9; Page 98; 524pp; English.
XX      CC
CC      Nucleotide sequences AAZ11972-212358 represent open reading frames (ORFs)
CC      of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
CC      proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
CC      their nucleic acids and antibodies are used for diagnosis, prevention (as
CC      vaccines) or treatment of Neisseria infections, such as meningitis,
CC      septicemia and gonorrhea. Both organisms are closely related. Fragments
CC      of the nucleic acids are useful as hybridisation probes and antisense
CC      reagents
XX      CC
SQ      Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,34e-157 Length: 963
Score: 1634.00 Matches: 317
Percent Similarity: 99.69% Conservative: 2
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 2 Gaps: 0

US-10-617-835-4 (1-320) x AAZ12006 (1-963)
Qy      1  MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db      1  ATGCAAGACGCGCGTGGATACCTATTTCTTTTCAGTTTATTTATTCGCGCTGGCGGG 60
Qy      21  ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGlnGluLeu 40
Db      61  ACATGACAGGTATTTCCATGCGATGCGAGGAGGTAAAGCTTTGGCGTGAACAAAGAACTT 120
Qy      41  ValAlaAlaSerAlaArgAlaAlaValLysAspMetCysLeuGlnAlaLeuHisGlyArg 60
Db      121  GTGGCGCTTCTTGCCAGAGCGTCCGTTAAAGACATGATTTTACAGGCATTAACAGAGAGA 180
Qy      61  LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerLysLeuThrGlyGly 80
Db      181  AAAGTTGATTTGATCATGTCCTACCTATGAGCGACCAAGGATTCAGGCAATTTGACAGGGGGT 240
Qy      81  ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100
Db      241  CGTACTCCATTTGATGACATGATTCGTGGCAATACATAAACAAGCCCTGCGTCCGTAAC 300
Qy      101  AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
Db      301  GATTACACCTATCCACGTTACGAAACCAACCGCTGAACAACATCAGCGGCTTGACAGGT 360
Qy      121  LeuThrThrSerLeuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db      361  TTAACCACTTTTATCTACTACTTAATGCCCCCTGCAGCTCTCGCACCAACCAACGAGT 420
Qy      141  SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160
Db      421  AGCGGAAGTAAAGCACTGCGGCTTAATAATTTGGCGGATGGGGATTTATCGAAATGAA 480
Qy      161  ThrLeuThrThrAsnProArgSerPheThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      481  ACCTTGACGACCTAACCCCGCGACACTGCTTTCTTTCCCACTTGTGTGACGCGTATTT 540
Qy      181  PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541  TTCCGCGCGGCGATAGACGTTGTTCTCTCTGCCAATGCCGATACAGATGTTATTTAAC 600
Qy      201  IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyAsnAlaGluThr 220
Db      601  ATCGACGATTTGGAAACGATACGCAACGAAACCGAAATGCACTTATACATGCCAACA 660
Qy      221  LeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysIleLeu 240

```

Db 661 CTGAAGCCCAACAACTGATATTTCCAGTACAGACAAATTAATAATTGCTC 720  
 Qy 241 ILeYsPProLYsThrAsnAlaPheGluAlaAlaYrLYsGluAsnTYrAlaLeuTrpMet 260  
 Db 721 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTTAAGAAATTAACGATTTGGATG 780  
 Qy 261 GLYProTYrLYsValSerLYsGlyLYsPProThrGluLYsLeuMetValAspPheSer 280  
 Db 781 GGGCCGTATAAGTAAGCAAGAAATTAACCGACGAAAGATTAATGTCGATTTCTCC 840  
 Qy 281 AspIleAsrProTYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 Db 841 GATATCCGACCATAGCGCATATCATCGGTAATCTCCGCCCATCCGTAGAGGCTGATAC 900  
 Qy 301 SerHisGluGlyTYrGlyTYrSerAspGluAlaValAlaArgGluHisArgGluGlnPro 320  
 Db 901 AGTCATGAGGGGTATGATGATACAGCATGATGAGTACGCAACAATAGACAAACACT 960  
 RESULT 9  
 AA253689  
 ID AA253689 standard; DNA, 963 BP.  
 XX  
 AC AA253689;  
 XX  
 DT 21-MAR-2000 (first entry)  
 XX  
 DE Neisseria meningitidis ORF 406 partial DNA sequence SEQ ID NO:1327.  
 XX  
 DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
 KW antibacterial; gene therapy; de.  
 XX  
 OS Neisseria meningitidis.  
 XX  
 PN WO957280-A2.  
 XX  
 PD 11-NOV-1999.  
 XX  
 XX 30-APR-1999; 99WO-US009346.  
 XX  
 PF 01-MAY-1998; 98US-0083758P.  
 PR 31-JUL-1998; 98US-0094869P.  
 PR 02-SEP-1998; 98US-0098994P.  
 PR 02-SEP-1998; 98US-0099062P.  
 PR 09-OCT-1998; 98US-0103749P.  
 PR 09-OCT-1998; 98US-0103794P.  
 PR 09-OCT-1998; 98US-0103796P.  
 PR 25-FEB-1999; 99US-0121528P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (GENO-) INST GENOMIC RES.  
 XX  
 P1 Frazer C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;  
 P1 Petersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarcella M;  
 P1 Tettelin H, Veneri JC;  
 PI  
 XX WPI: 2000-062150/05.  
 DR P-PSDB; AAY74927.  
 DR  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics.  
 XX  
 XX Claim 7, Page 728, 1453pp; English.  
 XX  
 CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of the  
 CC immunogenic compositions. The polypeptides can be used in the manufacture  
 CC of medicaments for treating or preventing infection due to Neisserial  
 CC bacteria (e.g. meningitis and septicemia), to detect the presence of

CC Neisseria bacteria, or to raise antibodies. They may also be used to  
 CC screen for agonists or antagonists, which may themselves have use as  
 CC antibacterial agents. The polynucleotides of the invention may also be  
 CC used in gene therapy protocols  
 SQ Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other:  
 Alignment Scores:  
 Pred. No.: 1,346-157 Length: 963  
 Score: 1634.00 Matches: 317  
 Percent Similarity: 99.69% Conservative: 2  
 Best Local Similarity: 99.06% Mismatches: 1  
 Query Match: 99.33% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-617-835-4 (1-320) x AA253689 (1-963)  
 Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 Db 1 ATGCAGCAGCGGCTCTGATACCTATTTCTTTTTCAGTTTATTTATTCGCCCTGGCGG 60  
 Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyLYsArgPheAlaValGluGlnLeu 40  
 Db 61 ACACGACAGGATTCATCGCATGCGGAGGTAAACGCTTGGCGTGAACAAGACTT 120  
 Qy 41 ValAlaAlaSerAlaArgAlaAlaValLYsAspMetAspLeuGluAlaLeuHisGlyArg 60  
 Db 121 GTGGCCGCTCTGCGCAGAGCTGCCCTTAAGCAAGATTTACAGCATTTACAGCAGCA 180  
 Qy 61 LysValAlaLeuTYrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
 Db 181 AAGTTGCAATTGTACATTGCCACTTGTGGCGGACCAAGTTTCAGGAGTTTGACAGGGGT 240  
 Qy 81 ArgTYrSerIleAspAlaLeuIleArgLYsGlyTYrIleAsnSerProAlaValArgThr 100  
 Db 241 CGCTACTCATTTGATGACAGCTGATTCGTGCGAATCATTAACAGCCCTGCCGCTGACC 300  
 Qy 101 AspTYrThrTYrProArgTYrGluThrThrAlaGluThrThrSerGlyLYsLeuThrGly 120  
 Db 301 GATTACACCTTATCCACGTTACGAAACCAACCGCTGAACACATCAAGGGGTTGACAGGT 360  
 Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 Db 361 TTAAACACTTCTTTATCTACACTTAATACCTGACCTCTCGACCAATACAGCGGT 420  
 Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTYrArgAnglu 160  
 Db 421 AGCGGAAGTAAGAAGCAGTCTGGGCTTAATAATTTGCGGGATGCGGGATTATGAAATGAA 480  
 Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 Db 481 ACCTTGAGACTTAACCCCGGACACCTGCTTTCTTCCACTTGTTACAGCCGATATT 540  
 Qy 181 PheLeuArgGlyIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 Db 541 TTCCTGGCGGCAATGACGCTTCTCTCTGCCAATGCCGATACGATGTGTTATTAC 600  
 Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTYrAsnAlaGluThr 220  
 Db 601 ATCGACGTAATCGGAACGATACGCAACGAAACCGAAATGACCTATACAAATGCCGAAACA 660  
 Qy 221 LeuLYsAlaGlnThrLYsLeuGluTYrPheAlaValAspArgThrAsnLYsLeuLeu 240  
 Db 661 CTGAAGCCCAACCAACCTGGAATATTTCCAGATGACAGAACCAATTAATAATTGCTC 720  
 Qy 241 ILeYsPProLYsThrAsnAlaPheGluAlaAlaYrLYsGluAsnTYrAlaLeuTrpMet 260  
 Db 721 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTTAAGAAATTAACGATTTGGATG 780  
 Qy 261 GLYProTYrLYsValSerLYsGlyLYsPProThrGluLYsLeuMetValAspPheSer 280  
 Db 781 GGGCCGTATAAGTAAGCAAGAAATTAACCGACGAAAGATTAATGTCGATTTCTCC 840

QY 281 AsplleatpProYrGlyAsnHietHrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
Db 841 GATATCCGACCATACGGCAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATAC 900  
QY 301 SerHisGluGlyTYrGlyTYrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320  
Db 901 AGTCATAGGGGTATGATACAGCGATGAAGTAGTCAGACAACATAGACAAGGACAACT 960  
RESULT 10  
ID AA254613 standard; DNA; 963 BP.  
XX AA254613;  
AC  
XX  
XX AA254613;  
AC  
XX  
XX 21-MAR-2000 (first entry)  
DT  
XX Neisseria meningitidis ORF 406 partial DNA sequence SEQ ID NO:3109.  
DE  
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KM antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy; ds.  
XX  
XX Neisseria meningitidis.  
OS  
XX  
XX MO9957280-A2.  
PN  
XX  
XX 11-NOV-1999.  
PD  
XX  
XX 30-APR-1999; 99WO-US0009346.  
PF  
XX  
XX 01-MAY-1998; 98US-0083758P.  
PR 31-JUL-1998; 98US-0094869P.  
PR 02-SEP-1998; 98US-0098994P.  
PR 02-SEP-1998; 98US-0099062P.  
PR 09-OCT-1998; 98US-0103749P.  
PR 09-OCT-1998; 98US-0103794P.  
PR 09-OCT-1998; 98US-0103796P.  
PR 25-FEB-1999; 99US-0121528P.  
XX  
XX  
XX (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX  
XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalati E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX  
XX WPI: 2000-062150/05.  
DR P-PSDB; AAY75811.  
XX  
XX  
XX Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.  
XX  
XX  
XX Example 1; Page 108-109; 1453pp; English.  
XX  
XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides  
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254617 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of the  
CC invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the manufacture  
CC of medicaments for treating or preventing infection due to Neisserial  
CC bacteria (e.g. meningitis and septicemia), to detect the presence of  
CC Neisseria bacteria, or to raise antibodies. They may also be used to  
CC screen for agonists or antagonists, which may themselves have used as  
CC antibacterial agents. The polynucleotides of the invention may also be  
CC used in gene therapy protocols  
XX  
XX Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;  
SQ  
Alignment Scores: 1.34e-157 Length: 963  
Pred. No.: 1634.00 Matches: 317  
Score:

Percent Similarity: 99.69% Conservative: 2  
Best Local Similarity: 99.06% Mismatches: 1  
Query Match: 99.33% Indels: 0  
Gaps: 0  
US-10-617-835-4 (1-320) x AA254613 (1-963)  
QY 1 MetArgAlaArgLeuLeuLeuProIleuPhePheSerValPheIleLeuSerAlaCysGly 20  
Db 1 ATGCAGACGACGGCTGCTGATACCTTATCTTTTCAGTTTATTTATTCGCCCTCGCGG 60  
QY 21 ThrLeuThrglyIleProSerHisGlyGlyGlyLysArgPheAlaValGlnGlnLeu 40  
Db 61 ACACGACAGAGGATTCATCGATCGATGCGGAGGTAAACGCTTGGCGTGAACAAGACTT 120  
QY 41 ValAlaAlaSerAlaArgAlaAlaValLysPheMetAspLeuGlnAlaLeuHisGlyArg 60  
Db 121 GTGGCCGCTCTGTGCAGAGCTGCGCTTAAAGACATGATTTACAGGCACTTACAGGACGA 180  
QY 61 LysValAlaLeuTYrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
Db 181 AAAGTTGATTTGATCATTTGCCACATATGGGCGACCAAGGTTCAAGGCAAGTTTACAGGGGGT 240  
QY 81 ArgTYrSerIleAspAlaLeuIleArgGlyGlyTYrIleAsnSerProAlaValArgThr 100  
Db 241 CGCTAATCTCATTTGATGACGATGATTCGTGCGCAATATCAACACCCCTGCGCTCGTACC 300  
QY 101 AspTYrThrTYrProArgTYrGlyLeuThrThralaGlnThrThrSerGlyGlyLeuThrGly 120  
Db 301 GATTACACCTTATTCACAGCTTACGAAACACCGCTTAAACACATACAGCGGCTTTACAGGT 360  
QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
Db 361 TTAACCACTTCTTATATCATCACTTAAGCCCTGACACTCTCTCCACCAATCAGACGGT 420  
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTYrArgAsnGln 160  
Db 421 AGCGAAGTAAACACAGCTGGGCTTAAATTTGCGGGAAGGGGATTAACGAAATGAA 480  
QY 161 ThrLeuThrThrAspProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
Db 481 ACCTTGAAGACTTAAACCGCGCGACACCTGCTTTTCCACTTGGTACAGACCGTATTT 540  
QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
Db 541 TTCCTGGCGGCAATAGACGTTGTTCTCTGCAATGCCATACAGATGTGTTATTATAC 600  
QY 201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHisLeuTYrAsnAlaGlnThr 220  
Db 601 ATCGACGTATTCGGAACGATACGACAAACCAATGACCTTATCAATGCCGAACA 660  
QY 221 LeuLysAlaGlnThrLysLeuGlyTYrPheAlaValAspArgThrAsnLysLeuLeu 240  
Db 661 CTGAAGCCCAAAACAAACCTGGAATATTTGCCATGACAGAACCAATAAAATTTGCTC 720  
QY 241 IleLysProIyThrAsnAlaPheGluAlaAlaTYrLysGluAsnTYrAlaLeuTYrMet 260  
Db 721 ATCAAAACCAAAAACCAATGCGTTTGAAGCTGCCATTAAGAAATTTGCGCTTGTGATG 780  
QY 261 GlyProTYrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer 280  
Db 781 GGGGCGTATTAAGTAACCAAAAGGAATTAACCGAGGAAGATTAAGTGTGATTTCTCC 840  
QY 281 AsplleatpProYrGlyAsnHietHrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
Db 841 GATATCCGACCATACGGCAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATAC 900  
QY 301 SerHisGluGlyTYrGlyTYrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320  
Db 901 AGTCATAGGGGTATGATACAGCGATGAAGTAGTCAGACAACATAGACAAGGACAACT 960  
RESULT 11  
AAA81296



ID AAA01296 standard; DNA; 963 BP.  
 XX AAA01296;  
 AC  
 XX 04-DEC-2000 (first entry)  
 DT  
 XX  
 XX N. meningitidis partial DNA sequence m406.seq seq ID NO:1034.  
 DE  
 XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
 KM antigen; vaccine; diagnosis; infection; antibacterial; identification;  
 KW Meningococcus B; MenB; da.  
 XX  
 XX Neisseria meningitidis.  
 OS  
 XX  
 XX MO200022430-A2.  
 XX  
 XX 20-APR-2000.  
 PD  
 XX  
 XX 08-OCT-1999; 99MO-US023573.  
 XX  
 XX 09-OCT-1998; 98US-0103794P.  
 PR  
 XX 30-APR-1999; 99US-0132068P.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 XX  
 XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Scarlato V,  
 PI Massignani V, Galeotti C, Mora M, Ratti G, Scarpelli M, Scarlato V,  
 PI Rapuoli R, Pizzi M;  
 XX  
 XX WPI: 2000-318079/27.  
 DR  
 XX P-PsDB; AAB25656.  
 XX  
 PT Isolated nucleotide sequences of Neisseria meningitidis which can be used  
 PT in the diagnosis and treatment of N. meningitidis infection and other  
 PT Neisserial infections, for example, N.gonorrhoea.  
 XX  
 PS Class 9; Page 107; 1760pp; English.  
 XX  
 XX The present invention describes methods of obtaining immunogenic proteins  
 CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent  
 CC specifically claimed Neisseria meningitidis genomic DNA sequences;  
 CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA  
 CC sequences and their corresponding proteins; AAA81254 to AAA81259 and  
 CC AAA81304 to AAA81321 represent PCR primers used in the isolation of  
 CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent  
 CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all  
 CC used in the exemplification of the present invention. The nucleic acid  
 CC sequences, protein sequences, and antibodies against them, can be used in  
 CC the manufacture of a composition. The composition can be used as a  
 CC medicament (or in the manufacture of a medicament) for treating,  
 CC preventing or diagnosing infection due to Neisserial bacteria. For  
 CC example, some of the identified proteins could be components of vaccines  
 CC against Meningococcus B; against all serotypes; and/or against all  
 CC pathogenic Neisseriae. Identification of sequences from the bacterium  
 CC will also facilitate production of biological probes, particularly  
 CC organism-specific probes. Attempts to make efficacious Meningococcus B  
 CC vaccines have failed mainly due to antigen tolerance. Multivalent  
 CC vaccines have also been tried but none have successfully overcome  
 CC antigenic variability. The provision of further, complete sequences may  
 CC provide an opportunity to identify secreted or surface exposed proteins  
 CC that may be presumed targets for the immune system and which are not  
 CC antigenically variable or at least more conserved than other more  
 CC variable regions  
 XX  
 XX Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1 34e-157 Length: 963  
 Score: 1634.00 Matches: 317  
 Percent Similarity: 99.69% Conservative: 2  
 Best Local Similarity: 99.06% Mismatches: 1  
 Query Match: 99.33% Indels: 0  
 DB: 3 Gaps: 0

US-10-617-835-4 (1-320) x AAA01296 (1-963)  
 QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 DB 1 ATGCAAGCAGCGCTGCTGATACCTATTCTTTTTCAGTTTATTATTATCCGCTCGGG 60  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyGlyValArgPheAlaValGluGluGlu 40  
 DB 61 ACACGACAGGATTCATTCATGATGCGGAGGTAAACGCTTTGCGGTCAACAGAACTT 120  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValIleValSerMetAspLeuGlnAlaLeuHisGlyArg 60  
 DB 121 GTGGCCGCTTCGCGACAGCTCCGTTAAAGACATGATTTACAGGCATTACAGGACCA 180  
 QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
 DB 181 AAAGTTGCATTTGATCATTTGCCACTATGGCGACCAAGGTTCAAGGAGTTGACAGGGGGT 240  
 QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
 DB 241 CGCTACTCCATTGATGACCTGATTCGTGGCGAATACATTAACAGCCCTGCGCTAC 300  
 QY 101 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120  
 DB 301 GATTACACCTATCCACGTTACGAAACACCGCTGAACACATCAAGCGGTTGACAGGT 360  
 QY 121 LeuThrThrSerLeuSerThrLeuAlaAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 DB 361 TTAAACCACTTCTTATCTACACTTAATGCCCCCTGCACTCTCGCACCAATACAGCGGT 420  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAlaIleGlyGlyMetGlyAspTyrArgGlu 160  
 DB 421 ACCGGAAGAAAGAGAGCTGGCTTAATATGCGGGATGAGGATTAATGAAATGA 480  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 DB 481 ACCTTGACGACATCAACCGCGGACACTGCTTTCTTCCACTGTGTACAGACCGATTT 540  
 QY 181 PheLeuArgGlyIleAspValValSerProAlaAlaAlaAspThrAspValPheIleAla 200  
 DB 541 TTCCTGCGGCGATAGACGTTGTTCTCTCGCAATGCGGATACAGATGTATTATTAC 600  
 QY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220  
 DB 601 ATCGACGATTTGGAAACGATAGCAACAGAACCTGAATGACCTATACAAATCCGAAACA 660  
 QY 221 LeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLysLeuLeu 240  
 DB 661 CTGAAGGCCCAACAACTGGAATTTTCGACATGACAGAACCAATAAATAATTGCTC 720  
 QY 241 IleLysProLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTyrMet 260  
 DB 721 ATCAAAACCAAAACCAATGCGTTGAGCTGCTTAAAGAAATATACCATTTGTGAGAG 780  
 QY 261 GlyProTyrLysValSerIleGlyLysProThrGluGlyLeuMetValAspPheSer 280  
 DB 781 GGGCGGTAAAGTAAGCAAGAAATTAACCGAGGAAGATTAATGATTCATTTCTC 840  
 QY 281 AspIleAspProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 DB 841 GATATCCAGCACTACGCGCATATACGCGTAATCTCGCCCATTCGTAAGGCTGATTAAC 900  
 QY 301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValAlaGlnHisArgGlnGlyGlnPro 320  
 DB 901 AGTCATGAGGGGTATGATATACAGCATGAGTATGCGCAACATATGACAAAGACAACT 960  
 RESULT 12  
 AAF21581  
 ID AAF21581 standard; DNA; 963 BP.  
 AC  
 XX AAF21581;  
 XX

DT 13-MAR-2001 (first entry)  
XX  
XX N. meningitidis partial DNA sequence m406.seq SEQ ID NO:74.  
XX  
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
KM ds.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO200066791-A1.  
XX  
PD 09-NOV-2000.  
XX  
PF 08-MAR-2000; 2000WO-US005928.  
XX  
PR 30-APR-1999; 99US-013206BP.  
XX PR 08-OCT-1999; 99MO-US023573.  
XX PR 28-FEB-2000; 2000GB-00004695.  
XX  
PA (CHIR) CHIRON CORP.  
XX PA (GENO) INST GENOMIC RES.  
XX  
PI Pizza M., Hickey B., Peterson J., Tetrelin H., Venter JC;  
PI Maestroni V., Galeotti C., Mora M., Ratti G., Scarselli M., Scarlato V;  
PI Rapunio R., Frazer CM., Grandi G;  
XX  
DR WPI; 2000-647603/62.  
XX P-PSDB; AAB58586.  
XX  
PT Neisseria meningitidis B full length genome sequence and open reading  
PT frames are used to detect, treat and prevent Neisserial infections.  
XX  
XX Example 1; Page 108; 692pp; English.

The present invention describes the full length genome of Neisseria meningitidis B (NMB). The sequences in AAP21544 and AAP21607 to AAP21613 represent fragments of the NMB genomic sequence, as the sequence was too long to go in a record on its own it was split into 8 sequences which overlap each other at the beginning and end of each sequence by 49980 bp (i.e. the last 49980 bp of AAP21544 is repeated at the beginning of AAP21607, and so on), AAP21545 to AAP21607 are repeated at the beginning of AAP21608, the last one), AAP21545 to AAP21588 encode the Neisseria proteins given in AAB58550 to AAB58593, and AAP21589 to AAP21606 represent PCR primers which are used in the exemplification of the present invention. The NMB genome and fragments from it have antibacterial activity, and can be used in vaccines and gene therapy. Neisseria nucleic acids, proteins and/or antibodies which binds to the proteins can be used in compositions for treating or preventing infection due to Neisserial bacteria or as a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised to Neisserial bacteria. Computers, computer memory, computer storage medium or computer databases can be used in a search to identify open reading frames (ORFs) or coding sequences within the NMB genome. The DNA sequences provide further opportunities to find antigenic or immunogenic proteins which are more effective in vaccines than the outer membrane proteins currently used

XX  
XX Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	
Score: 1.34e-157	Matches: 963	
	Mismatches: 317	
Percent Similarity: 99.69%	Conservative: 2	
Best Local Similarity: 99.06%	Mismatches: 1	
Query Match: 99.33%	Indels: 0	
	Gaps: 0	

US-10-617-835-4 (1-320) x AAP21581 (1-963)

D5 1 MetaGALIAAGTGTGTTGGTGAATTCATTCTTTTGAGTTTAAATTATTCGCCGCGGG 20  
1 ATGCACGACCGCCTGCTGTAACCTAATCTTTTTCAGTTTAAATTATTCGCCGCGGG 60

QY	21	ThrLeuThrnglylleProSerHiscllygilylylsarGPhaAlaValclugngluleu	40
Db	61	ACACTGACAGGATTTCATTCGCATGCGAGGTAAACGTTTGCGGTCCAAACAAGAACTT	120
QY	41	ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisgilyArg	60
Db	121	GTGGCGCGCTTCGCCAGAGCTGCCGTTAAAGACATGGATTTCAGAGCATTAACGGACGA	180
QY	61	LysValAlaLeuTyrlleAlaThrmecGlyAspGnglySerGlySerleuThrgily	80
Db	181	AAAGTTGACATTGTATGATGCCCTATGGCGCACCAAGTTTCAGGACGTTTGACAGGGGCT	240
QY	81	ArgTySerSerleAspAlaLeuIleArggilyguTyrlleAsnSerProAlaValArgThr	100
Db	241	CGCTACTCCATTGATGCACTGATTCGTGGCGAAATACATAACGCCCTCCGCTCCATAC	300
QY	101	AspTyrrThTyrrProArgTyrglyuThrrAlaGluThrrThrSerGilylyleuThrgly	120
Db	301	GATTACACCTATCCAGCTTACGAACCAACCGCTGAACAACTACAGGCGGTTTGACAGGT	360
QY	121	LeuThrTrhSerleuSerThrleuasnAlaProAlaLeuSerArgThrglnSerAspGly	140
Db	361	TTAACCACTCTTTATTCACCTTAATGCTCCCTGCACTCTCCGACCCCAATCAGCGGT	420
QY	141	SerGlySerArgSerSerleuGilyleuasnIleGilylymetGlyAspTyrrArgasnGlu	160
Db	421	AGCGGAAGTAAAGCAGCTGGGCTTTAAATATGGCGGATGGGGGATTCGAATGAATGA	480
QY	161	ThrLeuThrThrrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrrValPhe	180
Db	481	ACCTTGAAGACATAACCCCGCGACACTGCTTTCTTCCACTTGGTGTACAGACCGTATTT	540
QY	181	PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrrAspValPheIleAsn	200
Db	541	TTCCGCGCGGATGACGCTGTTCTCTCCCAATGCCGATRCAGATGCTTTATTAAAC	600
QY	201	IleAspValPheGlyThrrIleArgAsnArgThrrGluMetHisLeuTyrrAsnAlaGluThr	220
Db	601	ATCGACGTATTGCGAAACATACGCAACAGAACCGAAATGCACCTATACAAATCCCGAAACA	660
QY	221	LeuIysAlaGlnThrThyleuGluThyrPheAlaValaAspArgThrrAsnIysIleLeu	240
Db	661	CTGAAGGCCCAACAAACATCGAATATTTCCGACATGACAGAACCAATATAAAAAATTGCTC	720
QY	241	IleIysProIysThrrAsnAlaPheGluAlaAlaTyrrLyGluAsnTyrrAlaLeuTrpMet	260
Db	721	ATCAAAACCAAAAAACAATGCGTTTGAAGCTGCCCTTTAAAGAAATTAACCATTTGTCGAAG	780
QY	261	GlyProTyrrLysValSerIleGlyIleIysProThrrGnglyleuMetValAspPheSer	280
Db	781	GGGCGGTTAAAGTAAGCAAGAAATTAAACGACAGGAAGATTAATGATTCATTTCTCC	840
QY	281	AspIleArgProTyrrGlyAsnHisThrrGlyAsnSerAlaProSerValGluAlaAspAsn	300
Db	841	GATATCCGACCATACGCGCATCATATACGGGTAACTCCGCCCATCCGTAAAGCTGATTAAC	900
QY	301	SerHisGnglyTyrrGlyTyrrSerAspGluAlaValaArgGlnHisArgGnglyGlnPro	320
Db	901	AGTCATGAGGGGTATGATGATACAGCATGAAGTAGTCGCAACATAGACAAAGACAAACCT	960
RESULT 13			
AAZ12007			
ID	AAZ12007 standard; DNA, 963 BP.		
XX	AAZ12007;		
AC	08-OCT-1999 (first entry)		
DT			
XX	Neisseria meningitidis strain A complete ORF15 sequence.		
DE			
XX	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;		
KW	treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.		
XX			



XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen U, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX  
XX MPI: 2000-062150/05.  
DR P-PSDB; AAY75813.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.  
XX  
XX Example 1, Page 110; 1453pd; English.  
XX  
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of the  
CC invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the manufacture  
CC of medicaments for treating or preventing infection due to *Neisseria*  
CC bacteria (e.g. meningitis and septicemia), to detect the presence of  
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to  
CC screen for agonists or antagonists, which may themselves have use as  
CC antibacterial agents. The polynucleotides of the invention may also be  
CC used in gene therapy protocols  
XX  
SQ Sequence 963 BP, 285 A, 233 C, 218 G, 227 T, 0 U, 0 Other;

## Alignment Scores:

Pred. No.: 1,42e-156 Length: 963  
Score: 1624.00 Matches: 315  
Percent Similarity: 99.69% Conservative: 4  
Best Local Similarity: 98.44% Mismatches: 1  
Query Match: 98.72% Indels: 0  
DB: Gaps: 0

US-10-617-835-4 (1-320) x AAZ54615 (1-963)

QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCyGly 20  
DB 1 ATGGAGACGCGCGTGTATCTATCTTTTTCAGTTTATTTATTCGCGCTGCGG 60  
QY 21 ThrLeuThrglyIleProSerHISGlyGlyIleYsArgPheAlaValIGlnGlnIleu 40  
DB 61 ACACGTACAGGTATTCATCGCATGCGGAGGTAAACCTTCGCGGTGGAACAAGACTT 120  
QY 41 ValAlaSerAlaArgAlaAlaValIleYsAspMetAspLeuGlnAlaLeuHISGlyArg 60  
DB 121 GTGGCCGCTTCGCCAGAGCTGCCGTTAAAGACATGATTACAGGCACTTACAGGACGA 180  
QY 61 LysValAlaLeuYrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrgly 80  
DB 181 AAAGTTGATGTGATTCATTGCAACATAGGCGACCAAGGTTGACGATTCACAGGGGCT 240  
QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyIleYsIleAsnSerProAlaValArgThr 100  
DB 241 CGCTACTCCATTGATGCACTATTCGTGGCAATACAAACAGCCCTGCGTCCGTAAC 300  
QY 101 AspTyrThrYrProArgTyrGlyIleThrAlaGlnThrThrsArgGlyIleuThrgly 120  
DB 301 GATTACACCTATTCACGTTACGAACACACCGCTGAAACAAACATGAGCGGTTGACAGGT 360  
QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
DB 361 TTAAACCACTTTTATCTTACACTTAATGCCCTGACACTCTGCGGACCAATTCAGACGCT 420  
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
DB 421 AGCGGAAGTAAACAGAGCTGGCTTAAATTTTGGCGGGAAGGGGAGTATTCGAATGAA 480  
QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHISLeuValGlnThrValPhe 180

DB 481 ACCTTGAAGACTAACCCGCGGACACTGCTTCTTCCCACTTGATACAGACGTAATTT 540  
QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
DB 541 TTCCTGGCGGACATAGAGTGTTCCTGCGCAATGCCGATAGGATGTTTATTTAAC 600  
QY 201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHISLeuYrAsnAlaGlyThr 220  
DB 601 ATCAGCATTTTGGAAAGCATACGCAACGAAACCAATGCACCTTACAAATGCCGAACA 660  
QY 221 LeuLysAlaGlnThrIleYsLeuGlyIleThrPheAlaValAspArgThrAsnLysLeuLeu 240  
DB 661 CTGAAGGCCCAAAACAAATCGAATATTTCGCACTTACAGCAACCAATTAATAATTCCTC 720  
QY 241 IleYsProLysThrAsnAlaPheGlnAlaAlaTyrIleYsGlnAsnYrAlaLeuThrMet 260  
DB 721 ATCAACCAAAACCAATGCGTTTGAAGCTGCTTAAAGAAATTCAGCATTTGGATG 780  
QY 261 GlyProTyrIleYsValSerLysGlyIleYsProThrGlnGlyLeuMetValAspPheSer 280  
DB 781 GGACGTRTAAAGTAAACCAAGAAATTAACCGACAGAGATTAAATGTCGATTTCTCC 840  
QY 281 AspIleArgProTyrGlyAsnHISThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
DB 841 GATATCCAAACCATACGCGCAATCATATGCTAACTCTGCCATCCGTAGAGGCTGATAAC 900  
QY 301 SerHISGlnGlyIleYrGlyIleYrSerAspGluAlaValArgGlnHISArgGlnGlyPro 320  
DB 901 AGTCATAGGGGATGTGATTCAGCGATGAAGAGTGGACGACATTAACAAAGGCAACT 960  
RESULT 15  
AAZ53690  
ID AAZ53690 standard; DNA, 963 BP.  
XX  
AC AAZ53690;  
XX  
DT 21-MAR-2000 (first entry)  
XX  
DE *Neisseria meningitidis* ORF 406 partial DNA sequence SEQ ID NO:1329.  
XX  
KW *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy; ds.  
XX  
OS *Neisseria meningitidis*.  
XX  
EN W0957280-A2.  
XX  
PD 11-NOV-1999.  
XX  
PF 30-APR-1999; 99WO-US009346.  
XX  
PR 01-MAY-1998; 98US-0083758P.  
PR 31-JUL-1998; 98US-0094869P.  
PR 02-SEP-1998; 98US-0098994P.  
PR 02-SEP-1998; 98US-0099062P.  
PR 09-OCT-1998; 98US-0103749P.  
PR 09-OCT-1998; 98US-0103794P.  
PR 09-OCT-1998; 98US-0103796P.  
PR 25-FEB-1999; 99US-0121528P.  
XX  
PA (CHIRON) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen U, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX  
XX MPI: 2000-062150/05.  
DR P-PSDB; AAY74928.  
XX  
PT Novel *Neisseria* polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.

XX Claim 7, Page 729-730; 1453p; English.

CC AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AA275941  
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254673 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of the  
 CC invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the manufacture  
 CC of medicaments for treating or preventing infection due to *Neisseria*  
 CC bacteria (e.g., meningitis and septicemia), to detect the presence of  
 CC *Neisseria* bacteria, or to raise antibodies. They may also be used to  
 CC screen for agonists or antagonists, which may themselves have use as  
 CC antibacterial agents. The polynucleotides of the invention may also be  
 CC used in gene therapy protocols

XX Sequence 963 BP; 285 A; 233 C; 218 G; 227 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,42e-156 Length: 963  
 Score: 1624.00 Matches: 315  
 Percent Similarity: 99.69% Conservative: 4  
 Best Local Similarity: 98.44% Mismatches: 1  
 Query Match: 98.72% Indels: 0  
 DB: 3 Gaps: 0

US-10-617-835-4 (1-320) x AA253690 (1-963)

QY 1 MetPArgAlaArgLeuLeuIleProIleuPheSerValPheIleLeuSerAlaCysGly 20  
 DB 1 ATGCAAGCAGCGGCTGCTGATTAATCTTTTCAGTTTATTATTAATCCGCTGCGGG 60  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGlnGluLeu 40  
 DB 61 ACACGACAGAGGATTCATCCATCGATGCGGAGGTAAACGCTTCGCGTCAACAAGACTT 120  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 DB 121 GTGGCCGCTCTGCGCAGAGCTGCCGTAAAGCATGAGATTACAGGCACTTACACGACGA 180  
 QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
 DB 181 AAAGTTCATTGATGACATTAATGAGGCGACCAAGGTTCAAGCACTTTGACAGGGGCT 240  
 QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
 DB 241 CGCTACTCATTTGATGACACTGATTCGTGGCGAATACATAACAGCCCTGCCGTCCGTACC 300  
 QY 101 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120  
 DB 301 GATTACACCTATCCAGCTTACGAAACCAACCGCTGAACACATCAGCGGTTTGACAGGT 360  
 QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 DB 361 TTAAACCACTCTTATTAATCACTTAATGCCCCCTGCACTTCGCGCACCCAATCAGACCGT 420  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
 DB 421 AGCGGAATTAAGAAGAGCTGTGGCTTAATATATGCGGGATGGGGATTATCGAAATGAA 480  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 DB 481 ACCTTGACGACTTAACCCCGCGACACTGCTTTCTTCCCACTTGTTACAGCCGTATTT 540  
 QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 DB 541 TTCTCGCGGGATAGAGCTGTCTCTCGCAATGCCGATACGAGATGTTATTAAAC 600  
 QY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220  
 DB 601 ATCGACGATTTCGAAAGATACGAAACGAAACGAAATGCACTTATACATGCCGAAACA 660

QY 221 LeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLeuLeu 240  
 DB 661 CTGAAGCCCAACAACTGGAATATTTCGACGTAGACAGAACCAATTAATAATTGCTTC 720  
 QY 241 IleLysProLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuIleTrpMet 260  
 DB 721 ATCAACCAAAADACCAATGCGCTTGAAGCTCCCTATTAAGAAATTAACGATTGGATG 780  
 QY 261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSer 280  
 DB 781 GGACCGTATTAAGTAAAGCAAGGATTTAAACCGACAGAAAGATTAAATGTGATTTCTCC 840  
 QY 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 DB 841 GATATCCAAACCATACGGCAATCATATGGGTAATCTGCCCCATCCGTAGAGGCTGATAAC 900  
 QY 301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320  
 DB 901 AGTCATGAGGGGATGATGATACAGCGATGAACAGACGATGACATTAAGAGGCACT 960

Search completed: August 19, 2005, 01:04:08  
 Job time : 546 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 18, 2005, 22:53:39 ; Search time 3122 Seconds

(without alignment) 3901.527 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645

Sequence: 1 MRARLLRILPILSVFLLSACG.....SHEGYSDEAVRHRGQCP 320

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USP10.spool.p/US10617835/runat.18082005.115608.8376/app_query.fasta_1.519.
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10617835@cgn2.1.5180@runat.18082005.115608.8376 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : EST:\*

```
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g981:*
9: gb_g982:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	6.3	750	5	BP182164
2	103.5	6.3	557	4	BM819564 K-EST0087
3	101	6.1	724	8	BH515144 BQGS41TF
4	100.5	6.1	853	7	CF683309 CCAB21TF
5	100.5	6.1	882	7	CF687616 CCAC089TF
6	100.5	6.1	898	7	CF701318 CCAR073TO
7	100.5	6.1	2651	9	AY412499 Mus muscu
8	100.5	6.1	4036	3	AK049496 Mus muscu
9	97.5	5.9	638	4	BI543281 S1C_G08 S

C 10	97.5	5.9	808	7	CF720552	CCAH127TF
C 11	97.5	5.9	875	7	CF224113	AGENCOURT
C 12	97	5.9	1074	6	CA066594	SCBOARD101
C 13	96	5.8	800	8	B2517711	BOMRT19TF
C 14	95.5	5.8	553	6	CA192949	SCRLSB104
C 15	94.5	5.7	568	8	BH887997	LB01788A.
C 16	94.5	5.7	631	1	AL961261	AL629935
C 17	94.5	5.7	661	1	AL629935	AL629935
C 18	94.5	5.7	902	7	CN510835	AGENCOURT
C 19	94	5.7	688	5	BQ578598	WHE0307_C
C 20	94	5.7	813	7	CK200430	FGAS00894
C 21	94	5.7	1128	9	CL977717	OsIFCC030
C 22	93.5	5.7	862	7	CN209233	Tor9747 G
C 23	93.5	5.7	1165	6	CD508628	CD508628
C 24	93.5	5.7	1349	8	CL648116	CH213-162
C 25	93.5	5.7	1388	8	B2576008	m8t2-4730
C 26	93	5.7	581	5	BP255245	BP255245
C 27	93	5.7	797	7	CN509047	AGENCOURT
C 28	93	5.7	884	7	CN502363	AGENCOURT
C 29	93	5.7	932	7	CN510402	AGENCOURT
C 30	92.5	5.6	669	7	CF684102	CCAGY30TF
C 31	92.5	5.6	801	6	CB904901	LC1C040XG
C 32	92.5	5.6	801	7	CF876129	LC1C040XG
C 33	92.5	5.6	847	7	CK352302	AGENCOURT
C 34	92.5	5.6	879	7	CN507579	AGENCOURT
C 35	92.5	5.6	1052	3	CNS0958N	single re
C 36	92.5	5.6	1098	9	CNS05EXB	single re
C 37	92.5	5.6	1206	9	CL965795	OsIFCC012
C 38	92.5	5.6	491	2	BE106299	UI-R-B01
C 39	92	5.6	491	4	BI471703	ae82f0.
C 40	92	5.6	677	3	CNS099AK	single re
C 41	92	5.6	751	5	BX925187	BX925187
C 42	92	5.6	827	5	BU105657	603004222
C 43	92	5.6	831	6	CA487870	AGENCOURT
C 44	92	5.6	833	2	AW982176	HVSM9000
C 45	92	5.6				

#### ALIGNMENTS

RESULT 1  
BP182164 750 bp mRNA linear EST 14-OCT-2003  
LOCUS BP182164 NRPG Bombyx mori cDNA clone NRPG0202, mRNA sequence.

DEFINITION BP182164  
ACCESSION BP182164.1 GI:37662855

VERSION BP182164.1  
KEYWORDS  
SOURCE Bombyx mori (domestic silkworm)

ORGANISM Bombyx mori

REFERENCE 1 (bases 1 to 750)

AUTHORS Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.

TITLE Bombyx mori cDNA

JOURNAL Unpublished (2000)

COMMENT Contact: Mita K

National Institute of Agrobiological Sciences

Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan

Email: kmits@niae.affrc.go.jp

method: uni-directional, sequence direction: sequenced from 3' primer

(5' -> 3')

FEATURES

source

1..750

/organism="Bombyx mori"

/mol\_type="mRNA"

/strain="Skuko x Ryuhaku"

/db\_xref="taxon:7091"

/clone="NRPG0202"

/tissue\_type="pheromone gland"

/dev\_stage="adult"

/clone\_lib="NRPG"









```

QY 190 ProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleArgAsn 209
Db 449 CGAGCTCCTGGT-----TCATTTGAGCGCTTGGTTTCAGACAGCCCAT 408
QY 210 ArgThrGluMetHisIleuTyArgAsnIleGluThrIleuValAGlnThrIysLeuGluTyr 229
Db 407 CCGCGG-----GTTGCCGACACCAATGAGAACGCT 378
QY 230 PheAlaValAspArgThrAsnIlyValLeuIleValProIyThrAsnIlePheGlu 249
Db 377 TCTGCTGTCACTTCATTGACCTTACTGTTATCCCTTCAACGCGCGCT----- 324
QY 250 AlaAlaTyArgIleAsnTyArgIleuTyrMetGlyProTyArgValSerIyGlyIle 269
Db 323 TCAGCTTCCCGATTCACAGACAGTT-----GAGGTG 291
QY 270 LysProThrGluGlyLeuMetValAspPheSerAspIleArgProTyArgIlyAsnIleThr 289
Db 290 CGAGGAATGGATGCGCTTGTAGGGGATTTTACAGACTAATTGTAGCGGACAGGAGCG 231
QY 290 GlyAsnSerAlaProSerValGluIleAspAsnSer 301
Db 230 ATAAACCTCTCTCCAAAGTTGCAACGACGACAGTAC 195
RESULT 6
CF701318/c 898 bp mRNA linear EST 16-AUG-2004
LOCUS CCAFO73TO C.neoformans strain JEC21 Cryptococcus neoformans var.
ACCESSION CF701318
VERSION CF701318.1 GI:41555477
KEYWORDS EST
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
ORGANISM Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
REFERENCE 1 (bases 1 to 898)
AUTHORS Loftus,B.
TITLE End sequencing of clones from a full length enriched, normalized
JEC21 cDNA library
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: CCAFO73TR
Contact: Brendan Loftus
TIGR
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: crypt@tigr.org
Seq primer: TF.
FEATURES
source
1..898
Location/Qualifiers
/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"
/strain="JEC21"
/db_xref="taxon:40410"
/clone="CCAFO73"
/clone_lib="C.neoformans strain JEC21"
/notes="Vector: PCMVSPORT6; Site 1: Noti EcoRV. The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jenniffer Lodge"
ORIGIN
Alignment Scores:
Pred. No.: 1.25 Length: 898
Score: 100.50 Matches: 57
Percent Similarity: 35.328 Conservative: 32
Best Local Similarity: 22.624 Mismatches: 88
Query Match: 6.11% Indels: 75
DB: 7 Gaps: 10

```

```

US-10-617-835-4 (1-320) x CF701318 (1-898)
QY 90 GlyIuTyIleAsnSerProAlaValArgThrAspTyThrTyProArgTyGluThr 109
Db 821 GCGGCTATCGTAATATATCCCTCTGCTGCTCCGATCT-----GCTTATGAGGCG 768
QY 110 ThrIleGluThrThrSerGlyGlyLeuThr----- 119
Db 767 AACGCTGCTTACTTCGAGCACTATCTCCGCCCACTGAGTCCGACGGATGTATGCG 708
QY 120 -----GlyLeuThrThrSerLeuSer----- 126
Db 707 CAACTGAACTCGGAAGAAATTGGCTCCGGTTTGGAGATCTCACAAATTGGGGCTAGC 648
QY 127 -----ThrLeuAsnAlaProAlaLeu 133
Db 647 TACAGTCAAGGCTGCACGCGAGTGCAGGATGACCATGACATGCGGCTAGC 588
QY 134 SerArgThrGlnSerAspGlySerIySerArgSerSerLeuGlyLeuAsnIleGly---- 152
Db 587 TTGAGAACTCAGACGACACAGGAGACCTCGCGCTGATGTCGGACCGAAATCTGGTGT 528
QY 153 -----GlyMetGlyAspTyArgAsnGluThrIleuThrThrAsnProArgAspThr 169
Db 527 CCTGCTCGGCTGAGAGTTCTTCAAGAGGAGAGCTTTGCA-----CCATTGATGCG 474
QY 170 AlaPheLeuSerHisIleValGlnThrValPhePheLeuArgGlyIleAspValIleSer 189
Db 473 GCGGCTATCAATCCTCTTCCAAAGCTGTC-----GGCATTCAGATGCTAGAGC 426
QY 190 ProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleArgAsn 209
Db 425 CGAGCTCCTGGT-----TCATTTGAGCGCTTGGTTTCAGACAGCCCAT 384
QY 210 ArgThrGluMetHisIleuTyArgAsnIleGluThrIleuValAGlnThrIysLeuGluTyr 229
Db 383 CCGCGG-----GTTGCCGACACCAATGAGAACGCT 354
QY 220 PheAlaValAspArgThrAsnIlyValLeuIleValProIyThrAsnIlePheGlu 249
Db 353 TCTGCTGTCACTTCATTGACCTTACTGTTATCCCTTCAACGCGCGCT----- 300
QY 250 AlaAlaTyArgIleAsnTyArgIleuTyrMetGlyProTyArgValSerIyGlyIle 269
Db 299 TCAGCTTCCCGATTCACAGACAGT-----GAGGTG 267
QY 270 LysProThrGluGlyLeuMetValAspPheSerAspIleArgProTyArgIlyAsnIleThr 289
Db 266 CGAGGAATGGATGCGCTTGTAGGGGATTTTACAGACTAATTGTAGCGGACAGGAGCG 207
QY 290 GlyAsnSerAlaProSerValGluIleAspAsnSer 301
Db 206 ATAAACCTCTCTCCAAAGTTGCAACGACGACAGTAC 171
RESULT 7
AY412499 2651 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM4577 gene, VIRTUAL TRANSCRIPT, partial sequence.
ACCESSION AY412499
VERSION AY412499.1 GI:39768464
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2651)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.D., Slnsky,J.C.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous

```

gene tries  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 REFERENCE PUBMED 14671302  
 AUTHORS 2 (bases 1 to 2651)  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1..2651  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>2651  
 /locus\_tag="HGM4577"  
 gene  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 6.33 Length: 2651  
 Score: 100.50 Matches: 50  
 Percent Similarity: 39.00% Conservative: 28  
 Best Local Similarity: 25.00% Mismatches: 63  
 Query Match: 6.11% Indels: 59  
 DB: 9 Gaps: 11  
 US-10-617-835-4 (1-320) x AY412499 (1-2651)  
 Oy 76 SerleuThrglygl-----ArgTySerlleAspAlaleu 87  
 Db 1413 ACTCTGACTTCGGGAAACACACACCGAGGACCGTACCGGTACCTGATCCACC 1472  
 Oy 88 lle-----ArgglglutyrilleasnsrProalaValArgthAspTyThrTyPro 105  
 Db 1473 GTATCTCTCGGACAGCTACCGGAAAGAACCCCTTGAGAGCTCGAAGTAT----- 1523  
 Oy 106 ArgTyrgluthrThrlgluthrThrseryglyleuThrglyleuThrThrsleu 125  
 Db 1524 -----TCTCGAAGGCTGGTAC-----CTC 1544  
 Oy 126 SerThrlleuanaProalaIleuserArgThrglnsArpAglySeryglSeraRser 145  
 Db 1545 AGCGACGAGACTCACCCGAAGTCTG--AGAACTGATCTTACGACATCGATCGAGAC 1601  
 Oy 146 Serleugly-----leuanslle 151  
 Db 1602 AAACCTTGGGAAAGGAGAGAAACTGTCCCAAGAGTGTAGCAAAATGAATGACATC 1661  
 Oy 152 gllyglmetglAspTyArgAsnlglnthrlleuthrThrranProArgAspThrlalhe 171  
 Db 1662 GGTGCTC-----TTCAGACATTACAGCTTTGGGATCAACCCAAATGCTCGAGTAC 1712  
 Oy 172 leuserHlsleuValglnthrValPhepheleuArglylleAspValIalSerProla 191  
 Db 1713 ATATCCGGCCCTCATGAGGCTACACTTAT-----GAGCTGAGGCTTTAAACCGGCC 1766  
 Oy 192 AsnAlaAspThr--AspValPheIleasnIleasPvalPheglYThrIleArgAsnArg 210  
 Db 1767 CGGATAGATCGAAAGATGTGTTTGTCCCATTTCAGTAGACTCATGAAACAAAGCCGCA 1826  
 Oy 211 ThrGluMet-----HisleuTyAsnAlaIle 219  
 Db 1827 ACGGCTGTGCTACCTGTGCGACGACCTTTTACATGACGACCAACCTTCAACCTAAG 1886  
 Oy 220 ThrleuysalaglN---ThrlyleuaglutyrrPheAlaValAspArgThrslnlyls 238  
 Db 1887 ATTGAAGACGGCGACATCTGAAGCTGGTGGTGTTCACACTGGAGCCACCGCCGAGAGG 1946

AK049496  
 LOCUS 4036 bp mRNA linear HTC 03-APR-2004  
 DEFINITION Mus musculus 7 days embryo whole body cDNA, RIKEN full-length  
 enriched library, clone: C430017H16 product: weakly similar to  
 THRO1000381 PROTEIN (FRAGMENT) [Homo sapiens], full insert  
 sequence.  
 ACCESSION AK049496  
 VERSION AK049496.1 GI:26340229  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 11042159  
 2  
 3 Shibata, K., Itoh, M., Aizawa, K., Nagata, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Kitesuna, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishie, T., Harada, A.,  
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 11076861  
 4  
 5 The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 6 (bases 1 to 4036)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
 Hori, F., Imochi, K., Ishii, Y., Itoh, M., Kagawa, T., Kasuoka, T.,  
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.  
 Direct Submission  
 TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome

COMMENT

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse cDNAs.

Please visit our web site for further details.  
URL: <http://genome.gsc.riken.jp/>  
URL: <http://phantom.gsc.riken.jp/>.

## FEATURES

source

1..4036  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM DB:C430017H16"  
/db\_xref="taxon:10090"  
/clone="C430017H16"  
/tissue\_type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="7 days embryo"  
1..3292  
/note="unnamed protein product; putative  
weakly similar to THYROID00381 PROTEIN (FRAGMENT) [Homo  
sapiens] (SEPR|Q9H8A6, evidence: FASTY, 65.2%id,  
99.4%length, match=1605)"  
/codon\_start=2  
/protein\_id="BAC33778.1"  
/db\_xref="GI:26340230"  
/translation="SPRTQLSSCSRLBSLRNRPPPLQRMASDSCIRCGARDEDP  
OPRGGMGMSGSHRAATISALLPICHDPSGSSGRPKQDVGTGVSISSEDHPED  
LPKPPAVTVKKLQKMYKGRLLSLGMRGRDGTAPKPGAGTTSPLNLSGMVHNHY  
LSVTDORITLTDLFENVYGLPKKREEDVPALRGHPLANNITVSKKNMLYQS  
TLRSQSLDESKRSODRFRYSLSPVSPKHWVRHRCCHAVCNARATLSRNGSSAF  
EDNADDEGETWNRIPEDPSLGIAHVSLEPANNALKTLPVSMLSARDIMKAPPE  
DRGDIQTSOSNCINSADSMHSTELNQHCKQRIQHTQSPADSDPVLKASITGCGII  
VSTNTELRAMESSPSPPSKGSSITWSLPDKLSPPYVAKLSMKKRLPEFRKL  
GARGALHYVNNPDSSTPLSKMNCRIPOGVTLISGNTNMYISRLHDTTVSSRSYR  
KKPLGSSKYSCKGYLSDGSDPELRTRSKSGEHLKGRRTVSSCKMELDIGAF  
RHYSPADPKCKOYISGLMSVHFGAEGIKPRINDSKOVFCALQVDSVNAKRTALLTC  
RTTFLMDHTFRIETENAOHLKLVYFSMEPTRRRRVCCHGVTVLPAFRKTHQQLA  
VLEPRGLIYVATLMEOWENSLHGLDKRREAVMGVDIQVKEKENGAPLPLIK  
CIVEIEKRCQVGLYRLCGSAVAKELREAEKOSKTGVCENYPOINIVTGLKQ  
LYRLSPPLITKQLYEAVLDMAKSPLNKSSGCEPDSRLTVDLDCJPDYKAT  
LKLMLDLHLKLVASHYEVNMTCONLAVCFGPVLINROEASTHNRVFTDSBELASAL  
DFKHIIEVLHVLQMLPQVORLTVKREPSLCLTEGSSSLNLYRKKERCCVNLSSG  
SGVLRTRQARLDSPLSNRYACDMSGCSYSILCTGNCNKINDYDPLEGESRYSK  
TAGAEVMTAQRIPMSGGCTFTYTLTVITESTIVDQANLRLQESITDILIGLERELN  
KNKLMSV"

## CDS

polyA\_signal  
4009..4014  
/note="putative"  
polyA\_site  
4036  
/note="putative"  
ORIGIN  
/note="putative"

## Alignment Scores:

Pred. No.: 11.9 Length: 4036  
Score: 100.50 Matches: 50  
Percent Similarity: 39.00% Conservative: 28  
Beet Local Similarity: 25.00% Mismatches: 63  
Query Match: 6.11% Indels: 59  
DB: 3 Gaps: 11

US-10-617-835-4 (1-320) x AK049496 (1-4036)

QY 76 SerLeuThrGlyGly-----ArgTYSerIleAspAlaLeu 87  
DB 1442 ACTGACTCTCCGGGAAACCAACGAGACGTGATAGCCGCTACCTTATACAC 1501  
QY 88 Ile-----ArgGlyGlyIleAsnSerProAlaValArgThrAspTYrThrYrPro 105  
DB 1502 GATCTCTCCGGACACGCTACCGAAGAAACCCCTTGGAGCTCGAAGAT----- 1552  
QY 106 ArgTYrGlyThrAlaGlyThrSerGlyGlyLeuThrGlyLeuThrThrSerIleu 125  
DB 1553 -----TCCTGCAAGGCTGGGTAC-----CTC 1573

QY 126 SerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerArgGlySerArgSer 145  
DB 1574 AGCGACGAGACTACCCGAACTG---AGAACTAGATCTAGCAAGCATGATCGAGAC 1630  
QY 146 SerLeuGly-----LeuAsnIle 151  
DB 1631 AAACCTTGGGAAAGGAGAGAAACTGTCCCAAGCACTTACCAAAATGAACTAGAC 1690  
QY 152 GlyGlyMetGlyAspTYrArgAsnGlyThrLeuThrThrAsnProArgAspThrIaPhe 171  
DB 1691 GGTGCC-----TTGACACATPACAGCTTTCGCGATCAACCAATGCTCGCAGTAC 1741  
QY 172 LeuSerHisIleValGlnThrValPhePheLeuArgGlyIleAspValSerProAla 191  
DB 1742 ATATCCGCTCTTACAGCGTACACTCTAT-----GGAGCTGAGGTTTAAACCGGCC 1795  
QY 192 AsnAlaAspThr---AspValPheIleAsnIleAspValPheGlyThrIleArgAsnArg 210  
DB 1796 CGGATGATTTGAAAGATGTGTTTGTCCATTCAGTACGTAAGTCAACAAAGCCAGA 1855  
QY 211 ThrGluMet-----HisLeuTYrAsnAlaGly 219  
DB 1856 ACGGCCCTGCTACCTGTGCGACGACACTTCTTAGACATGACACCAACGTTCAACATAGAG 1915  
QY 220 ThrLeuValAlaGln---ThrIleLeuGlyThrPheAlaValAspArgThrAsnIleAs 238  
DB 1916 ATTGAATAATGGCCAGCATCTGAAGCTGTGTGTTTCAGCTGGAGCCACCGCGAGAG 1975

RESULT 9  
LOCUS B1543281  
DEFINITION B1543281 638 bp mRNA linear EST 04-SEP-2001

SIC608 Sugar Beet stress germination cDNA library (subtracted)  
Beta vulgaris cDNA 5' similar to xyloglucan  
endo-1,4-beta-D-glucanase, mRNA sequence.

ACCESSION B1543281  
VERSION B1543281.1 GI:15427459

KEYWORDS EST, vulgaris (sugar beet)

SOURCE Beta vulgaris

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Caryophyllales; Amaranthaceae; Beta.

REFERENCE 1 (bases 1 to 638)  
de Joss Reyes, B.G., McGrath, J.M., Myers, S. and Derrico, C.

Differential gene expression in sugar beet seedlings (Beta

vulgaris) germinated under stress conditions

Unpublished (2000)

Contact: J. Mitchell McGrath

Sugar Beet Genetics Lab., USDA/ARS/Sugar Beet and Bean Research

Unit Michigan State University

494 Plant and Soil Sci. Bldg., East Lansing MI 48824, USA

Tel: 517 353 9262

Fax: (517)-337-6782

Email: [mitchcgm@msu.edu](mailto:mitchcgm@msu.edu)

Seq primer: T3.

FEATURES  
source  
1..638  
/organism="Beta vulgaris"  
/mol\_type="mRNA"  
/culivar="USH20"  
/db\_xref="taxon:161934"  
/tissue\_type="whole seedlings"  
/dev\_stage="4-day germination under stress (salt/NaCl,  
osmotic/Mannitol)"  
/lab\_host="XLI-Blue MRP"  
/clone\_lib="Sugar Beet stress germination cDNA library  
(subtracted)"  
/note="Organ: seeds; Vector: pBluescript II KS (+);  
Site\_1: EcoRI; Site\_2: XhoI; The mRNA samples used for  
this library were derived from seedlings germinated for 4  
days in 150mM NaCl and 200mM mannitol. The mRNA pool was

used as template for double stranded cDNA synthesis using the Stratagene Bluescript XR cDNA synthesis and library kit. The resulting cDNA was used as a tester for subtraction against a driver cDNA population derived from 4-day old seedlings germinated in moist filter paper. Subtraction was performed using the Invitrogen Photobiotin-Streptavidin subductor kit. The cDNA library was generated by directional ligation of the subcloned cDNAs in the EcoRI and XhoI sites of pBluescript XR vector (Stratagene). "

## ORIGIN

## Alignment Scores:

Pred. No.:	1.68	Length:	638
Score:	97.50	Matches:	48
Percent Similarity:	37.07%	Conservative:	28
Best Local Similarity:	23.41%	Mismatches:	68
Query Match:	5.93%	Indels:	61
DB:	4	Gaps:	10

US-10-617-835-4 (1-320) x B1543281 (1-638)

```

Qy 24 GYIleProSerH:sglyglylysaArgPheAlaVal]GlunGln------ 39
Db 23 GGACTTCATCAATGGGGGACATGAGGTGCTGTGTGACACCTCAGATTATTTATGTG 82
Qy 40 -----LeuValAlaAlaSerAlaArgAlaAla---ValLyAspMetAspLeuGln 55
Db 83 CTATATTCATGAGTGAGCATGACTGATGTCATGATGTCATTTTGACGATCGCTCCACAA 142
Qy 56 AlaleuHsIgaYgValAlaAlaLeuYrIlealathmetGlyAspGlnGlySerGly 75
Db 143 ACTTTTGAGGTGAGGGGTCAAAACATAT-----GGTTGGGGT 181
Qy 76 -----SerLeuThrGlyGlyArgTySerIleAspAlaLeuIleArgGlyGlu 91
Db 182 GGTGATCATCTCTGCTCTCTCATTAAGACCATGACTCT-----GGATCC 223
Qy 92 TyrIleAsnSerProAlaValArgThrAspYrThrTyProArgYrGluThrThraAla 111
Db 224 GGCCTTACCTCC-----AAAAAGACTTACTTATTGGAGATTGATATGCAATG 274
Qy 112 GluThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuAsnAlaPro 131
Db 275 AAACCTGCTGCTGTAACCTCGACGACCTGCTCATCTTTTATTTGCATCTGACCT 334
Qy 132 AlaleuSerArgThrGln-----SerAspGlySerGlySerArg 144
Db 335 AGCACTGGTAGACATGACGAGATCGACTTTGAGTTTAGAAATGCCAGTGACAGCCT 394
Qy 145 SerSerLeuGlyLeuAsnIleGlyGlyMetGly----- 155
Db 395 TACACCTTCACACCAACGATGTTGGCGAAGGAAAGGAGATGAGGACAGCAATTTCGA 454
Qy 156 -----AspYrArgAsnGluThrLeuThrThrAsnProArg--- 167
Db 455 CTCTGCTTGAACCAACCAATGAATACACACTTTCATTTGTTGGAACCCGAGATC 514
Qy 168 -----AspThr-----AlaPheLeuSerHisLeuValGlnThr 178
Db 515 ATCATGTTCTTGTCGATGAATAACACCAATTAAGATATTTTTAACCATGAGAAACACGGA 574
Qy 179 ValPhePheLeuArg 183
Db 575 GTTCCTTCTTAAG 589

```

RESULT 10  
 CP720552/c 808 bp mRNA linear EST 16-AUG-2004  
 LOCUS CCAHI27TF C.neoformans strain JEC21 Cryptococcus neoformans var.  
 DEFINITION CCAHI27TF C.neoformans strain JEC21 Cryptococcus neoformans var.  
 ACCESSION CP720552  
 VERSION CP720552.1 GI:41574711

## KEYWORDS

EST.

## SOURCE

Cryptococcus neoformans var. neoformans (Filobasidiella neoformans var. neoformans)

## ORGANISM

Cryptococcus neoformans var. neoformans

Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.

1 (bases 1 to 808)

Loftus, B.

End sequencing of clones from a full length enriched, normalized

JEC21 cDNA library

Unpublished (2003)

Other ESTs: CCAHI27TF

Contact: Brendan Loftus

TIGR

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: crypt@tigr.org

Seq primer: TF.

Location/Qualifiers

1..808

/organism="Cryptococcus neoformans var. neoformans"

/mol\_type="mRNA"

/strain="JEC21"

/db\_xref="taxon:40410"

/clone="CCAHI27"

/clone\_lib="C.neoformans strain JEC21"

/note="Vector: pCMVSPORT6; Site\_1: NotI EcoRV; The full length, normalized library was prepared from a variety of conditions using Jennifer Lodge"

## ORIGIN

## Alignment Scores:

Pred. No.:	2.39	Length:	808
Score:	97.50	Matches:	57
Percent Similarity:	34.92%	Conservative:	31
Best Local Similarity:	22.62%	Mismatches:	89
Query Match:	5.93%	Indels:	75
DB:	7	Gaps:	10

US-10-617-835-4 (1-320) x CP720552 (1-808)

```

Qy 90 GYIuTyrIleAsnSerProAlaValArgThrAspYrThrTyProArgYrGluThr 109
Db 783 GGCCTTATGTAATATATCCCTGCTGCTGCTGCCGATTT-----GCTTATGAGGC 730
Qy 110 ThrAlaGluThrThrSerGlyGlyLeuThr----- 119
Db 729 AAGCTGCTATACCTCGGACATATCCGCCACCTGATCCGACGAGATGTATGCG 670
Qy 120 -----GlyLeuThrThrSerLeuSer----- 126
Db 669 CACTGGAATCGGCAAAAGAAATGGCTCCGTTGGAGATTCAGCAATTGTGGCTTAC 610
Qy 127 -----ThrLeuAsnAlaProAlaLeu 133
Db 609 TACATCCAGGCTGCACGCGAGTGGGATATGCCAGATGACATGTCAGACCTTCAAGT 550
Qy 134 SerArgThrGlnSerAspGlySerGlySerArgSerLeuGlyLeuAsnIleGly--- 152
Db 549 TTGAGAGTCAAGATCAAGAGGAGACTCGGCTTCACTGTCGAGACGAAATCTTGCTGT 490
Qy 153 -----GlyMetGlyAspYrArgAsnGluThrLeuThrThrAsnProArgAspThr 169
Db 489 CCTGCTCGGCTGAGAGTTCTTCAAGAGGAGAGACTTTTGA-----CAATTGATGCG 436
Qy 170 AlaPheLeuSerHisLeuValGlnThrValPhePheLeuArgGlyIleAspValSer 189
Db 435 GCGCCTATTCATATCTTCCAAAGCTGTC-----GCCATTCCAGTGTGAGCG 388

```

QY 190 ProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleArgAsn 209  
 |||||  
 Db 387 CGAGTCTCTGGT-----TCAAATTGAGCGCTTGGTTGACACCCCAT 346  
 |||||  
 QY 210 ArgThrGluMetHisIleuTyraAsnAlaGluThrIleuTyraGluThrIleuTyra 229  
 |||||  
 Db 345 CCGCGG-----GTTGCCGACACCAATAGAACCGT 316  
 |||||  
 QY 220 PheAlaValAspArgThrAsnIleuTyraIleuTyraProIleuTyraAsnAlaPheGlu 249  
 |||||  
 Db 315 TCTGCTGCAACTATTAGCCCTAGTACTCTGTATCCCTTCAACGGGCGCT----- 262  
 |||||  
 QY 250 AlaAlaTyrrGluGluAsnTyrrAlaIleuTyrrMetGlyProTyrrIleValSerIleGlyIle 269  
 |||||  
 Db 261 TCAGCTTTGCCCGAATACACACGACGTT-----GAGGTG 229  
 |||||  
 QY 270 LysPrrThrGluGluIleuMetValAspPheSerAspIleArgProTyrrGlyAsnIleThr 289  
 |||||  
 Db 228 CGAGGAATGATGATCGCTTCTAGGGATTTTACACAGCTAATTGTAGCGGACAGGACG 169  
 |||||  
 QY 290 GlyAsnSerAlaProSerValGluAlaAspAsnSer 301  
 |||||  
 Db 168 ATAAACCTCTCTCAAGGTTGCAACGACGACGATGAC 133  
 |||||  
 RESULT 11  
 CF224113/c 875 bp mRNA linear EST 04-AUG-2003  
 LOCUS AGENCOURT\_15069986 NICHD\_XGC Emb7 xenopus tropicalis cDNA clone  
 DEFINITION IMAGE:6977012 5', mRNA sequence.  
 ACCESSION CF224113  
 VERSION CF224113.1 GI:33424821  
 KEYWORDS EST.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 Xenopodine; Xenopus; Silurana.  
 1 (bases 1 to 875)  
 NIH-MGC http://mgc.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NCI  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: c9apbs-r@mail.nih.gov  
 Tissue Procurement: Robert M. Grainger  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM14631 row: b column: 19  
 High quality sequence start: 22  
 High quality sequence stop: 56.  
 Location/Qualifiers  
 1..875  
 /organism="Xenopus tropicalis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8364"  
 /clone="IMAGE:6977012"  
 /cbase\_type="embryo, stages 20-27"  
 /dev\_stage="embryo, stages 20-27"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NICHD XGC Emb7"  
 /note="Vector: pCMV-SFOR6.1; Site 1: NotI; Site 2: EcoRV;  
 cloned unidirectionally. Primer: Oligo dT. Average insert  
 size 2.1 kb. Constructed by Invitrogen. Note: This is a  
 Xenopus Gene Collection (XGC) library."

ORIGIN  
 Alignment Scores:

Pred. No.: 2.7 Length: 875  
 Score: 57.50 Matches: 46  
 Percent Similarity: 39.88% Conservatve: 21  
 Best Local Similarity: 27.38% Mismatches: 70  
 Query Match: 5.93% Indels: 31  
 DB: 7 Gaps: 6  
 US-10-617-835-4 (1-320) x CF224113 (1-875)  
 QY 21 ThrIleuThrGlyIleProSerHisGlyGlyIleTyraArgPheAlaValGluGluIleu 40  
 |||||  
 Db 666 TCGTTCCTCATTGATTCGTGCCACGCT-----GAAAGGTGGGA 628  
 |||||  
 QY 41 ValAlaAsnAlaArgAlaValAlaValAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 |||||  
 Db 627 GTGGGTGCCAGGCGCCGCGTACGACCTCCCAATACGATCTGTGTAAGCAGCTTCMAAGA 568  
 |||||  
 QY 61 LysValAlaLeuTyrrIleAlaThrMetGlyAspGlnGlySer-----GlySerIleuThr 78  
 |||||  
 Db 567 AGTGTGCTCTTTCACAGAGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 508  
 |||||  
 QY 79 GlyIleArgTyrrSerIleAspAlaLeuIleArgGlyIleTyrrIleAsnSerProAlaVal 98  
 |||||  
 Db 507 GCGGCACTGCAGAAAGCCAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 448  
 |||||  
 QY 99 ArgThrAspTyrrThrTyrrProArgTyrrGluThrAlaGluThrThrSerGlyGlyIleu 118  
 |||||  
 Db 447 CGGACTGAA-----GCGCAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 415  
 |||||  
 QY 119 ThrGlyIleuThrThrSerIleuSerThrIleuAsn-----AlaProAlaLeu 133  
 |||||  
 Db 414 CGAGTTAACCGCACTAGTGTGCTGACCTTCATAGGGGTGAAGAGGCTGTACACTTCTG 355  
 |||||  
 QY 134 SerArgThrGlnSerAspGlySerGlySerIleuGlyLeuAsnIleGlyIle 153  
 |||||  
 Db 354 -----CATGATCTGATCAGAGCTTGGGCAACCAATCTCGTGGC 316  
 |||||  
 QY 154 MetGlyAspTyrrArgAsnGluThrIleuThrThrIleuPro---ArgAspThrAlaPheIleu 172  
 |||||  
 Db 315 CGAGCAACGAGCGCATCTCCACCATCTTCGCGCTGCCAGGTGAGGTGTGACAGTCAT 256  
 |||||  
 QY 173 SerHisLeuValGlnThrValPhe 180  
 |||||  
 Db 255 TCTCATCTCTGACGTCCCACTTC 232  
 |||||  
 RESULT 12  
 CA066694 1074 bp mRNA linear EST 23-SEP-2003  
 LOCUS SCQAD01019F10.g AD1 Saccharum officinarum cDNA clone SCQAD01019F10  
 DEFINITION 5', mRNA sequence.  
 ACCESSION CA066694  
 VERSION CA066694.1 GI:34918218  
 KEYWORDS EST.  
 SOURCE Saccharum officinarum  
 ORGANISM Saccharum officinarum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
 complex.  
 1 (bases 1 to 1074)  
 Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.  
 The libraries that made SUCESF  
 Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
 Contact: Arruda P  
 Centro de Biologia Molecular e Engenharia Genetica  
 Universidade Estadual de Campinas  
 Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
 Tel: 55 19 3788 1137  
 Fax: 55 19 3788 1089  
 Email: parvuda@unicamp.br  
 Clone distribution: Clone distribution information can be found  
 through the Brazilian Clone Collection Center (BCCC) at  
 http://www.bcccenter.fcav.unesp.br

Plate: 019 row: F column: 10  
Seq primer: T7 Promoter Primer.

## FEATURES

SOURCE

Location/Qualifiers

1. 1074  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCRAD1019F10"  
/lab\_host="DH10B"  
/clone\_lib="AD1"  
/note="Organ: seedlings inoculated with Gluconacetobacter diazotrophicus; Vector: pSPori; Site\_1: SalI; Site\_2: NotI; An unidirectional cDNA library generated from (seedlings inoculated with Gluconacetobacter diazotrophicus). cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucrest.lad.ic.unicamp.br/public>"

## ORIGIN

## Alignment Scores:

Pred. No.:	4.19	Length:	1074
Score:	97.00	Matches:	37
Percent Similarity:	42.52%	Conservative:	17
Best Local Similarity:	29.13%	Mismatches:	61
Query Match:	5.90%	Indels:	12
DB:	6	Gaps:	4

US-10-617-835-4 (1-320) x CA066694 (1-1074)

Qy 32 LysArGpHeaIaValaGluGluGluValaAlaIaSerAlaArgAlaAlaValLysAsp 51  
Db 252 AAGGCTTCAGAGGCGAGCTGACGCTTGCTGCTCCGCGCGCGCTGTCAAG--- 308  
Qy 52 MetAspLeuGlnAlaLeuHisGlyYArgLysValAla--LeuTYrIleAlaThrMetGly 70  
Db 309 GTGAAGACCGAGTCATCGACGCGGAGGTGATCGCGGTAGAGCTGAGAGTGAAGGCC 368  
Qy 71 AspGlnGlySerGlySerLeuThrGlyGlyArgTYrSer----- 83  
Db 369 GAATCCGTCGACGAGGTGCTGGGGGAGCCAGATCGATGCTCTCCGCGCGAG 428  
Qy 84 IleAspAlaLeuIleArgGlyGluTYrIleAsnSerProAlaValArgThAspTYrThr 103  
Db 429 GCCACCGCAAGGTGAAGAGGAGCCGACCGATTCGCGCAAGAGGCCGACGATTC 488  
Qy 104 TYrProArgTYrGluThrThAlaGluThrThSerGlyGlyLeuThrGlyLeuThrThr 123  
Db 489 GTCCCCCGGCGCGCTGTGGGCGGCGCGGAGAACCGGGGGGCTCCCGGCTTAAAGCC 548  
Qy 124 SerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGlySer 143  
Db 549 CTTGATTC-----CCGCGGGGCTTGAAAGAAAGAGAGGCGCTTAAGGGGAG 599  
Qy 144 ArgSerSerLeuGlyLeuAsn 150  
Db 600 AGGAACCTTTGGAGAAAGAAC 620

## RESULT 13

BZ517711

LOCUS BZ517711 800 bp DNA linear GSS 16-DEC-2002  
DEFINITION BOMRT19TF BO\_2\_3\_KB Brassica oleracea genomic clone BOMRT19,  
genomic survey sequence.

ACCESSION BZ517711

VERSION BZ517711.1 GI:27048264

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 800)  
AUTHORS Town C.D., Van Aken S., Uterback T., Koo H. and Fraser C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL Unpublished (2001)  
COMMENT Other GSSs: BOMRT19TR  
Contact: Chris Town  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

## FEATURES

SOURCE

## ORIGIN

Pred. No.:	3.53	Length:	800
Score:	96.00	Matches:	58
Percent Similarity:	33.96%	Conservative:	33
Best Local Similarity:	21.64%	Mismatches:	99
Query Match:	5.84%	Indels:	78
DB:	8	Gaps:	12

US-10-617-835-4 (1-320) x BZ517711 (1-800)

Qy 41 ValAlaIaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeu----- 57  
Db 26 ATGACCGCTGAGATGATCGCTCAAGTCACAGACAAAGATCCGCGTTTCGTTGAAA 85  
Qy 58 -----HisGlyArgLysValAlaLeuTYrIleAlaThrMetGlyAspGln 72  
Db 86 AATATTATTTCCCGCCGCGGT-----TTACATTATGATGATGATCCGATTAAC 133  
Qy 73 GlySerGlySerLeuThrGlyGlyArgTYrSerIleAspAlaLeuIleArgGlyGluTYr 92  
Db 134 GATTCACAGTTCATGATGAGCAAT-----TCAAGAGATTC 169  
Qy 93 IleAsnSerProAlaValArgThAspTYrThTYrProArgTYrGluThrThAlaGlu 112  
Db 170 TGTGGCAATGGAACATTGACATTAAGCCCTCCACTCTCTGTTACCGCAAGT----- 223  
Qy 113 ThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuAsnAlaProAla 132  
Db 224 -----AATGCCAGGCA 235  
Qy 133 LeuSerArgThrGlnSerAspGlySerGlySerArgSerLeuGlyLeuAsnIleGly 152  
Db 236 GAATCTCCAAACAACATCATCATGATTAAGGCTCGATCGAATAAAGGCT 295  
Qy 153 GlyMetGlyAspTYrArgAsnGluThrLeu-----ThrThAsnProArgAspThr 169  
Db 296 CATGGGCTGACGAACATCGATGAGTCTGTGGAGCCATCGACACACCCGCGGAGATCG 355  
Qy 170 AlaIleLeuSerHisLeuValGlnThrValPhePheLeuArg---GlyIleAspVal 188  
Db 356 ACT-----AATCGACATTTTTCCTCCCTGCGCTTAAGGCGTGAAGAGCATG 400  
Qy 189 SerProAlaAsnAlaAspThrAsp-----ValPhe 198  
Db 401 GCTCTCTGTAAGTAACGTTTCAAGCTCCGACGTTCCAAATGCTCAATACGTGAG 460



REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
199	Ileana	-----IleapvalphegylthriilearganargThclumethis	214	
461	CTGAACAAAGACATCTCTCTGATATCCCTTATGTGATGAAGAAGACAGAGTTAACT	520		
215	LeuTryaenalaagluThrleuAlaaglInhTrlySleuGluTyPheAlaValaIsparg	234		
521	CTGCTGCACATCAAAATTATCAACATCATGATAGAGACTTACTAC-----	565		
225	ThraanlyblybLeuLeuIlelybPro-----Lyethraenala	247		
566	--AACCAAAAAGCTCTGCTGCCAGCATCTTGAACTCGGTGACTTCGTATGGCAAGTG	622		
248	PhegluAlaAlaTyTyLyseGluasnTy-----AlaLeuTryMetGlyProTy	263		
623	TTTGAAACACTTAAGAGCTTAACCCCGGTAACTCGGCGCAGGTGGAGACTTAC	682		
264	LySValserLySGlyIlelybPro	271		
683	AAATCATCAAAAGTGTCAACT	706		
RESULT 14				
CA192949	553 bp	mRNA	linear	EST 24-SEP-2003
LOCUS	SCRLSB1044F07.9	SBI Saccharum officinarum	CDNA clone	SCRLSB1044F07
DEFINITION	5', mRNA sequence.			
ACCESSION	CA192949			
VERSION	CA192949.1	GI:35139652		
KEYWORDS	EST.			
SOURCE	Saccharum officinarum			
ORGANISM	Saccharum officinarum			
REFERENCE	1 (bases 1 to 553)			
AUTHORS	Vecore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.			
TITLE	The libraries that made SUEST			
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)			
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: patricia@unicamp.br Clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br plate: 044 row: F column: 07 Seq primer: T7 Promoter Primer. Location/Qualifiers 1..553 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCRLSB1044F07" /lab_host="DH10B" /clone_1lb="SBI" /note="Organ: Stalk Bark from adult plants; Vector: pSport1, Site_1: SalI, Site_2: NotI. An unidirectional CDNA library generated from [Stalk Bark from adult plants]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose cl-2B 40cm-column and fragments sizing between 0.8 and 1.5 kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://bucest.fad.ic.unicamp.br/public"			

	Percent Similarity:	43.35%	Conservative:	25
	Best Local Similarity:	28.90%	Mismatch:	52
	Query Matchn:	5.81%	Indels:	46
DB:		6	Gaps:	10
US-10-617-835-4 (1-320) x CA192949 (1-553)				
OY	3	AAlargrleuenuileProlleleuPheserValpheilleuSerAiaCyB-----	19	
Dd	73	GGCGGCTACTCTCCGCCGTCTCCTCCTGCCTCCGCTCCGCGCGCGCGCGCGCTGCC	132	
OY	20	-----GlYThrleuthrGlyIleProSerHieGlyGly 31		
Dd	133	GGAGAGCTCGTGGAGAGACGGCGCATGGAGAATCTCCGCCCTCCCCCTCTCTCCCGC	192	
OY	32	LysArGrPhealVaIaGIuIngInluLeuVaIAlaIaseralAArgAlaIaVallYaAp	51	
Dd	193	CGGCGGTTCC-----CCCCTCCGCGCGCGCTGCAC	222	
OY	52	MetAspleuEngInAla-----LeuHieGlyArGlySValAlaLeuTYrlleAlaThmet	69	
Dd	223	CTG---ATCCGTGGCTGCMACTTCACC CGCGCGATGCTGCCGCCCTTCACCGCC	279	
OY	70	GIyAspIngIySerGIySerleuThrGIyGIyArGYrSerIleAspAlaLeuile-Ar	89	
Dd	280	GGCTGTAGAGCGCGCTCCGCCCGCGGCG-----ACCTGTGTGAG	321	
OY	89	ggIyGIuTyrrIleAsnerProAlaValaArgThrAspTYrThr---TYrProArGYrGI	108	
Dd	322	AGGCCCAATCGCGCTCGGCTCTTGCG---CGAGCCAGAGCGAGCGCGGACGTGGTAG	378	
OY	108	uThrThralaGIuThr-----ThrSerGIyGlyleuThrGlyleuThrThrSerleuSe	126	
Dd	379	GACCTCGGCCACCCAGCGGGGTACTACCGCTCCGCCAACCCAGCACGACGACGATGCTC	438	
OY	126	rThleuAsnAlaProAlaleuSerArGYrThGIuSerAsp-----	139	
Dd	439	TACTTTTTTCTTCTAG-----TCAGAGGGCGACAGACGACCCGTGGGAATCTGGCT	492	
OY	140	-----GlySerGIySerArGerSerleuGIyleu 149		
Dd	493	CACGCGCGCGCGCGGTGCAGAAACGAGCTCGGCTC	529	
RESULT 15				
LOCUS	BH887997	568 bp	DNA	linear GSS 07-AUG-2002
DEFINITION	LB01788a.d.SP6.1 Leishmania major Friedlin BAC Library Leishmania			
KEYWORDS	BH887997			
VERSION	BH887997.1	GI:22133688		
SOURCE	GSS.			
ORGANISM	Leishmania major			
REFERENCE	Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae;			
AUTHORS	Leishmania.			
JOURNAL	1 (bases 1 to 568)			
COMMENT	Myler,P.J., Vogt,C., Munden,H., Robertson,L., Siek,E., Faelinla,G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E., Stuart,K. and Ragland,M. Leishmania major Friedlin BAC End Sequences Unpublished (2002) Other_GSSs: LB01788a.d.T7.1 Contact: Myler PJ Seattle Biomedical Research Institute 4 Nickerson Street, Seattle, WA 98109-1651, USA Tel: 206 284-8846 Fax: 206 284-0313 Email: mylerpj@bri.org Seq primer: SP6 Class: BAC ends, Location/Qualifiers 1..568 /organism="Leishmania major"			

```
/mol_type="genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="LB01/88a"
/lab_host="E. coli GeneHogs + TrfA"
/clone_lib="Leishmania major Friedlin BAC library"
/notes="Vector: pCG270; Site 1: HindIII; Genomic DNA from
Leishmania major Friedlin in agarose blocks was partially
digested with HindIII, size selected, and ligated with
HindIII-digested pCG270 vector DNA. 10368 clones were
picked and arrayed in 384- and 96-well plates. Library
construction and arraying was carried out by ResGen
Corporation and clones and filters are available from
them"
```

## ORIGIN

## Alignment Scores:

Pred. No.:	3.16	Length:	568
Score:	94.50	Matches:	38
Percent Similarity:	37.50%	Conservative:	13
Best Local Similarity:	27.94%	Mismatches:	38
Query Match:	5.74%	Indels:	47
DB:	8	Gaps:	6

US-10-617-835-4 (1-320) x BH87997 (1-568)

```
QY 27 SerHisGlyGlyGlyArgPheAlaValGlnGlnLeuValAlaAlaSerAlaArg 46
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 154 GCTCAGGCGGGGTCGCCGTTTGTGGCAGAGAGGACCTTATTGCTCAACGGCG 213
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 47 AlaAlaValIysAspMetAspLeuGlnAlaLeuHisGlyArgIysValAlaLeuTyrIle 66
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 214 CAC-----CACGGCCGC-----CACGGCCGC-----CACGGCCGC----- 225
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 67 AlaThrMetGlyAspGln-----CACGGCCGC-----CACGGCCGC----- 72
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 226 ---ACCCCTTCAGAGCAGTCGGATGCGCTGCGCGCAACGTGCGGTCTGTTGCCGCT 282
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 73 GlySerGlySerLeuThrGlyGlyArgTyrSerIleAspAlaLeuIleArgGlyGluTyr 92
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 283 GGTGTCAGAACATGATGTTGGTGGCGCCACGCCGTGACTGCTTG----- 327
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 93 IleAsnSerProAlaValArgThrAspTyrThrTyrProArgTyrGluThrThrAlaGlu 112
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 328 -----CCGAGCTCGCGCGCC-----GTTACAACTTTGGAG 357
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 113 ThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuAsnAlaProAla 132
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 358 GCGGCGCAGAGGGGCGGCTTGGCAAACTATCCCGACAGCAGCACC---ACATGCAGTGGC 414
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 133 LeuSerArgThrGlnSerAspGlySerGlySerArgSerSerLeuGly 148
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 415 TCCTTCGCGACCCGACGCTGCGAATGTTGGCATCTTGACCTGCTTGGC 462
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: August 18, 2005, 23:49:11  
Job time : 3136 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_p2n model

Run on: August 19, 2005, 01:04:16 ; Search time 3941 Seconds  
(without alignments)  
3934.450 Million cell updates/sec

Title: US-10-617-835-4  
Sequence: 1 MKARLLIPILPSVFLSACG.....SHEGYGSDAVRHRGCGP 320

Scoring table:  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US10617835/runat.18082005.115730.8749/app.query.faeta\_1.519  
-DB=GenEmbl -QFMT=fasta -SUFFIX=p2n01lgo.rge -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIT=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10617835@cgn2.1.5600@runat.18082005.115730.8749 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	320	100.0	1136	6	A61825 Sequence 5
2	320	100.0	1136	6	AR393746 Sequence 5
3	320	100.0	3287	6	A61821 Sequence 1
4	320	100.0	3287	6	AR393745 Sequence 1

5	320	100.0	3294	1	AF142582	AF142582 Neisseria
6	282	88.1	963	6	A96050	A96050 Sequence 83
7	282	88.1	963	6	AX043997	AX043997 Sequence
8	207	64.7	963	6	A96046	A96046 Sequence 79
9	207	64.7	963	6	AX043995	AX043995 Sequence
10	194	60.6	10869	1	AE002559	AE002559 Neisseria
11	194	60.6	172325	6	AX044035	AX044035 Sequence
12	179	55.9	963	6	A96048	A96048 Sequence 81
13	179	55.9	963	6	AX043999	AX043999 Sequence
14	179	55.9	340806	1	NMA122491	AL162752 Neisseria
15	69	21.6	640	6	A96044	A96044 Sequence 77
16	21	6.6	942	6	A96280	A96280 Sequence 31
17	21	6.6	942	6	A96282	A96282 Sequence 31
18	21	6.6	942	6	A96284	A96284 Sequence 31
19	21	6.6	10057	1	AE002420	AE002420 Neisseria
20	21	6.6	11700	1	AE002393	AE002393 Neisseria
21	21	6.6	195767	1	NMA722491	AX162758 Neisseria
22	21	6.6	349980	6	AX044029	AX044029 Sequence
23	21	6.6	349980	6	AX044030	AX044030 Sequence
24	17	5.3	591	6	A96278	A96278 Sequence 31
25	10	3.1	156705	2	AC027270	AC027270 Homo sapi
26	10	3.1	166236	9	AC022832	AC022832 Homo sapi
27	10	3.1	167849	2	AC017001	AC017001 Homo sapi
28	10	3.1	170517	2	AC069526	AC069526 Homo sapi
29	10	3.1	175121	9	AC069246	AC069246 Homo sapi
30	10	3.1	176017	9	AP002748	AP002748 Homo sapi
31	10	3.1	182717	2	AC079823	AC079823 Homo sapi
32	10	3.1	193264	9	AC090885	AC090885 Homo sapi
33	10	3.1	198597	9	AC026218	AC026218 Homo sapi
34	10	3.1	202844	9	AC090956	AC090956 Homo sapi
35	10	3.1	229408	9	AC069271	AC069271 Homo sapi
36	10	3.1	231972	2	AC068055	AC068055 Homo sapi
37	9	2.8	663	8	AK107113	AK107113 Oryza sat
38	9	2.8	758	6	BD145166	BD145166 Primer fo
39	9	2.8	758	6	AX865104	AX865104 Sequence
40	9	2.8	919	9	SLY011914	AJ011914 Solanum l
41	9	2.8	2339	5	BC067995	BC067995 Xenopus t
42	9	2.8	10638	1	AE012264	AE012264 Xanthomon
43	9	2.8	12705	6	AX345051	AX345051 Sequence
44	9	2.8	12771	2	AC013927	AC013927 Drosophila
45	9	2.8	16720	6	AX251815	AX251815 Sequence

## ALIGNMENTS

RESULT 1	A61825	1136 bp	DNA	linear	PAT 09-MAR-1998
LOCUS	A61825	Sequence 5 from Patent WO9711181.			
DEFINITION	A61825				
ACCESSION	A61825.1	GI:3715996			
VERSION					
KEYWORDS					
SOURCE	Neisseria gonorrhoeae				
ORGANISM	Neisseria gonorrhoeae				
REFERENCE	1	Meyer, T.F., Rudel, T., Scheuerpflug, I., Fischer, Eckhard and			
AUTHORS	Maler, J.				
TITLE	NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE				
JOURNAL	ADHESION OF NEISSERIA CELLS TO HUMAN CELLS				
COMMENT	Patent: WO 9711181-A 5 27-MAR-1997;				
FEATURES	MAX PLANCK GESELLSCHAFT (DE)				
source	Other publication DE 19534579 970320.				
	location/Qualifiers				
	1..1136				
	/organism="Neisseria gonorrhoeae"				
	/mol_type="unassigned DNA"				
	/strain="MS11"				
	/db_xref="taxon:485"				

Alignment Scores:

Pred. No.: 0 Length: 1136  
 Score: 320.00 Matches: 320  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x A61825 (1-1136)

```

QY      1 MetAgaAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCyseGly 20
DB      135 ATGGGGGACGGCTGCTGATACCTATTCCTTTTGGTTTATTTATTCGCGCTGCGGG 194
QY      21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGluLeu 40
DB      195 ACACGACAGAGTATTCATCCGATGCGGAGGCAACCTTCGCGGTGGAACAGAACCTT 254
QY      41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGluAlaLeuHisGlyArg 60
DB      255 GTGGCCGCTTCGCGCAGAGCTGCCGTTAAAGCATGATTTACAGGCATTTACACGACGA 314
QY      61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
DB      315 AAAGTTGCATTGTACATTGCAACTATGGCGCACCAAGGTTCAAGGCAGTTGACAGGGGGT 374
QY      81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100
DB      375 CGCTACTCATTTGATGCACTGATTCGCGCGCAATACATTAACAGCCCTGCGCTCGCAC 434
QY      101 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyLeuThrGly 120
DB      435 GATTACACCTATCCGCGTTACGAAACCAACCGCTGAAACACATCAGCGGTTTGAACGAGT 494
QY      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB      495 TTACACCACTCTCTTATCTACACTTAATGCCCCCTGACCTCGCGCACTCAACATCAGAC 554
QY      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160
DB      555 AGCGGAATAGAGAGAGCTGGGCTTAATATTTGGCGGAGATGGGGATTATCGAAATATA 614
QY      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB      615 ACCTTGACAGACCAACCCCGCGACACTGCTTTCTTCCCACTGGTCAACAGCGTATTT 674
QY      181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
DB      675 TTCCGCGCGCATAGACGTTGTTCTCTCGCAATGCGCATACAGATGTGTTATTATAC 734
QY      201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220
DB      735 ATCGACGTAATTCGGAACGATACGCAACGAAACCGAAATGCACTTATACATGCGGAAACA 794
QY      221 LeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLeuLeu 240
DB      795 CTGAAAGCCCAACAACTGGAATATTTCCAGATAGACAGAACCAATTAATAATTTGCTC 854
QY      241 IleLysProLysThrAsnAlaPheGluAlaIleTyrLysGluAsnTyrAlaLeuTrpMet 260
DB      855 ATCAAAACCAAAACCAATGCGTTGAGCTGCTATTAAGAAATATGACATGTGGAGTG 914
QY      261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSer 280
DB      915 GGGCGGTAATAAGTAGCAAGAAAGAAATCAAAACGGAAGGATTATGTGATTTCTCC 974
QY      281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300
DB      975 GATATCCGGCCATACGGCAATCATACGGGTAACTCCGCCCATCGTGAAGCGTATATAC 1034
QY      301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGluPro 320
DB      1035 AGTCATGAGGGGATGATACAGCGATGAAGAGTGACAAACATTAAGCAAGGCAACCT 1094

```

RESULT 2  
 AR393746 LOCUS 1136 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 6 from patent US 6617128.  
 ACCESSION AR393746  
 VERSION AR393746.1 GI:40120650  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1136)  
 AUTHORS Meyer,T.F., Rudel,T., Schuenepflug,I., Maier,J., Eickernjäger,S.,  
 Schwan,T. and Fischer,B.  
 TITLE Nucleic acid molecules encoding proteins which impact the adhesion  
 of neisseria cells to human cells  
 JOURNAL Patent: US 6617128-A 6 09-SEP-2003;  
 FEATURES  
 source Location/Qualifiers  
 1. .1136  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ALIGNMENT Scores:

Pred. No.: 0 Length: 1136  
 Score: 320.00 Matches: 320  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x AR393746 (1-1136)

```

QY      1 MetAgaAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCyseGly 20
DB      135 ATGGGGGACGGCTGCTGATACCTATTCCTTTTGGTTTATTTATTCGCGCTGCGGG 194
QY      21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGluLeu 40
DB      195 ACACGACAGAGTATTCATCCGATGCGGAGGCAACGCTTCGCGGTGGAACAGAACCTT 254
QY      41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGluAlaLeuHisGlyArg 60
DB      255 GTGGCCGCTTCGCGCAGAGCTGCCGTTAAAGCATGATTTACAGGCATTTACACGACGA 314
QY      61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
DB      315 AAAGTTGCATTGTACATTGCAACTATGGCGCACCAAGGTTCAAGGCAGTTGACAGGGGGT 374
QY      81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100
DB      375 CGCTACTCATTTGATGCACTGATTCGCGCGCAATACATTAACAGCCCTGCGCTCGCAC 434
QY      101 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyLeuThrGly 120
DB      435 GATTACACCTATTCGGAACGATACGCAACCGCTGAAACCAATCAGCGGTTGACGAGT 494
QY      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB      495 TTAAACCACTCTTATTCATACCTTAATGCCCCCTGCACTCTCGCGCAACCATCAGACGT 554
QY      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160
DB      555 AGCGGAAGTAGAGACAGCTGGGCTTAATATTTGGCGGAGATGGGGATTATCGAAATATA 614
QY      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB      615 ACCTTGACAGACCAACCCCGCGACACTGCTTTCTTCCCACTGGTGAAGCCGATTTT 674
QY      181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
DB      675 TTCCGCGCGCATAGACGTTGTTCTCTCGCAATGCGCATACAGATGTGTTATTATAC 734
QY      201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220

```

Db 735 ATCGAGCTATTTCGAAACGATACGCAACAGAACCGAATACACTATACATGCCCCAACA 794  
 Qy 221 LeuLYSAlaGlnThrLYSLeuGlnTYrPheAlaValASPArgThrAsnLYSLeuLeu 240  
 Db 795 CTGAAGCCCAACAACTGGAATATTTCCGAGTAGACAGAACCAATTAATAATTGCTC 854  
 Qy 241 IleLYSProLYSThrAsnAlaPheGlnAlaAlaTYrLYSGlnAsnTYrAlaLeuTrpMet 260  
 Db 855 ATCAAAACCCAAACCAATCGTTTGAAGCTGCTTAATAAGAAATTAACCATGTCGAGAG 914  
 Qy 261 GLYProTYrLYSValSerLYSGlyIleLYSProThrGlnGlyLeuMetValASPheSer 280  
 Db 915 GGGCCGCTAATAAGTAAGCAAGAAATCAACCGACGAGATTAATGCTCATTTCTCC 974  
 Qy 281 AspIleArgProTYrGlyValAsnHISThrGlyValAsnSerAlaProSerValGlnAlaAspAsn 300  
 Db 975 GATATCCGGCCATACGGCAATCATACGGGTTAACTCCGCCCATCCGTAGAGGCTGATAAC 1034  
 Qy 301 SerHISGlnGlyTYrGlyTYrSerAspGlnAlaValArgGlnHISArgGlnGlyGlnPro 320  
 Db 1035 AGTCATGAGGGGTATGATACAGCATGAGCAAGTCGCAACATAGACAAAGGCAACT 1094

RESULT 3  
 A61821 3287 bp DNA linear PAT 09-MAR-1998  
 LOCUS Sequence 1 from Patent WO9711181.  
 DEFINITION A61821  
 ACCESSION A61821 GI:3715995  
 VERSION A61821.1  
 KEYWORDS  
 SOURCE Neisseria gonorrhoeae  
 ORGANISM Neisseria gonorrhoeae  
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 Neisseriaceae; Neisseria.

REFERENCE  
 AUTHORS Meyer,T.F., Rudel,T., Scheuerflug,I., Fischer, Eckhard and Maier,J.  
 TITLE NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE ADHESION OF NEISSERIA CELLS TO HUMAN CELLS  
 JOURNAL Patent: WO 9711181-A 1 27-MAR-1997;  
 MAX PLANCK GEBELTSCHAFT (DB)  
 COMMENT Other publication DE 19534579 970320.  
 FEATURES  
 source location/Qualifiers  
 1..3287  
 /organism="Neisseria gonorrhoeae"  
 /mol\_type="unassigned DNA"  
 /strain="MS11"  
 /db\_xref="taxon:485"  
 /clone\_lib="GENOMISCHE BIBLIOTHEK IN PBA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 3287  
 Score: 320.00 Matches: 320  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x A61821 (1-3287)

Qy 1 MetArgAlaArgLeuLeuLeuProIleLeuPheSerValPheIleLeuSerAlaCYGgly 20  
 Db 583 ATGGCGGCGCGGCTGTGATACCTATCTTTTCGTTTATTTATTCGCCCTGCGGG 642  
 Qy 21 ThrLeuThrGlyIleProSerHISGlyGlyLYSArgPheAlaValGlnGlnLeu 40  
 Db 643 ACACTGACAGATTCATCGCATGCGGAGCAACCGCTTGCGGTGAACAAGACTT 702  
 Qy 41 ValAlaAlaSerAlaArgAlaAlaValLYSAspMetAspLeuGlnAlaLeuHISGlyArg 60  
 Db 703 GTGGCGGCTTCTGCGAGACTGCCGTTTAAGAATGAGATTTCAGGCATTACACGACGA 762

Qy 61 LYsValAlaLeuTYrIleAlaThrMetGlyValAspGlnLYSLeuTYrGlyGly 80  
 Db 763 AAAGTTGATGTGTATGATTCAGCACTATGCGCGCAAGGTTCCAGGAGTTTGACAGGGGT 822  
 Qy 81 ArgTYSerIleASPAlaLeuIleArgGlyGlnTYrIleAsnSerProAlaValArgThr 100  
 Db 823 CGCTACTCATTTGATGACCTGATTTGGCGCGAATCATTAACAGCCCTGCGCTCGCAC 882  
 Qy 101 AspTYrTYrProArgTYrGlyThrAlaGlnTYrThrSerGlyLYSLeuThrGly 120  
 Db 883 GATTACACCTTATCCCGTTACGAAACCAACCGTGTAAACAAATCATGAGCGGTTTTCAGGGT 942  
 Qy 121 LeuThrTrsSerLeuSerThrLeuAsnAlaProAlaLeuSerArgTYrGlnSerAspGly 140  
 Db 943 TTACCACTTCTTTATCTACACTTAATGCCCTGCACTCTGCGCAACCAATCAGACGT 1002  
 Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTYrArgAnglu 160  
 Db 1003 AGCGAGATGAGAGCAGCTGGGCTTAATATTTGGCGGAGTGGGATTTATCGAAATGAA 1062  
 Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHISLeuValGlnThrValPhe 180  
 Db 1063 ACCTTGACGACCAACCCCGCGACACTGCTCTTCTCCACTTGTCACAGCCGATTT 1122  
 Qy 181 PheLeuArgGlyIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 Db 1123 TTCTGCGCGGCTATGACGCTTGTCTCTCTGCAATGCCATACAGATGTGTTATTAC 1182  
 Qy 201 IleAspValPheGlyTYrIleArgAsnArgThrGlnMetHISLeuTYrAsnAlaGlnThr 220  
 Db 1183 ATCGAGTATTCGGAACGATACGCAACGAAACCGAATGCACTTATACAAATGCCAACA 1242  
 Qy 221 LeuLYSAlaGlnThrLYSLeuGlnTYrPheAlaValASPArgThrAsnLYSLeuLeu 240  
 Db 1243 CTGAAGCCCAACAACTGGAATATTTCCAGATGACAGAACCAATTAATAATTGCTC 1302  
 Qy 241 IleLYSProLYSThrAsnAlaPheGlnAlaAlaTYrLYSGlnAsnTYrAlaLeuTrpMet 260  
 Db 1303 ATCAAAACCCAAACCAATGCGTTTGAAGCTGCTTAATAAGAAATTAAGCATTTGCGATG 1362  
 Qy 261 GLYProTYrLYSValSerLYSGlyIleLYSProThrGlnGlyLeuMetValASPheSer 280  
 Db 1363 GGGCGCTAATAAGTAAGCAAGAAATCAACCGACGAGAAATTAATGTCGATTTCTCC 1422  
 Qy 281 AspIleArgProTYrGlyValAsnHISThrGlyValAsnSerAlaProSerValGlnAlaAspAsn 300  
 Db 1423 GATATCCGGCCATACGGCAATCATACGGGTTAACTCCGCCCATCCGTAGAGGCTGATAAC 1482  
 Qy 301 SerHISGlnGlyTYrGlyTYrSerAspGlnAlaValArgGlnHISArgGlnGlyGlnPro 320  
 Db 1483 AGTCATGAGGGGTATGATACAGCATGACGATGACGACATGACAAATGACAAAGGCAACT 1542

RESULT 4  
 AR393745 3287 bp DNA linear PAT 18-DEC-2003  
 LOCUS Sequence 1 from patent US 6617128.  
 DEFINITION AR393745  
 ACCESSION AR393745  
 VERSION AR393745.1 GI:40120648  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE  
 AUTHORS Unclaseified.  
 1 (bases 1 to 3287)  
 Meyer,T.F., Rudel,T., Scheuerflug,I., Maier,J., Bickernjager,S., Schwan,T. and Fischer,E.

TITLE Nucleic acid molecules encoding proteins which impact the adhesion of neisseria cells to human cells  
 JOURNAL Patent: US 6617128-A 1 09-SEP-2003;  
 FEATURES location/Qualifiers

1..3287  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN

## Alignment Scores:

Pred. No.:	0	Length:	3287
Score:	320.00	Matches:	320
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-617-835-4 (1-320) x AK9393745 (1-3287)

```

Qy      1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db      583 ATGGCGGACGAGCTGCTGATACCTATCTTTTGGTTTATTTATTTATTCGCGCGCGG 642
Qy      21 ThrLeuThrGlyIleProSerHisGlyGlyIlyAspGpheaIaValGlnGlnLeu 40
Db      643 ACACGACAGAGTATTCATCCGATGCGGAGCGCAAAACCTTCGCGGTGAAACAAGACTT 702
Qy      41 ValAlaAlaSerAlaArgAlaAlaValIlyAspMetAspLeuGlnAlaLeuHisGlyArg 60
Db      703 GTGGCGCGTTCTGCAGAGCTGCCCTTAAAGACATGATTTACAGGCATTTACAGGACGA 762
Qy      61 LysValAlaLeuYrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Db      763 AAAGTTGCATTGTACATTGCAACTATGAGGCGAACAGGTTGACGACAGTTCACAGGGGGT 822
Qy      81 ArgGlySerIleAspAlaLeuIleArgGlyGlyIlyIleAsnSerProAlaValArgThr 100
Db      823 CGCTACTCATTGATGACCTGATTCGCGCAATACATTAACAGCCCTGCCGTCGCGACCC 882
Qy      101 AspYrThrYrProArgYrGlyIlyThrAlaGlnThrSerGlyGlyLeuThrGly 120
Db      883 GATTACACCTATCCGCGGTACGAACAACCCCTGAAACAACATCAGCGGCTTACAGCGGT 942
Qy      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db      943 TTAAACACTCTTTATCTACACTTAATGCCCTGCACCTCGCGACCCCAATCAGACCGGT 1002
Qy      141 SerGlySerArgSerSerLeuGlyLeuAsnGlnIleGlyGlyMetGlyAspYrArgAsnGly 160
Db      1003 AGCGGAATGAGAGAGCTGCTGGCTTAAATATTTGGCGGAGATGGGGATTATCGAATGAA 1062
Qy      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      1063 ACCTTGACAGCAACACCGCGGACACCTGCTTTCTTCCACTTGATACAGACCGTATTT 1122
Qy      181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      1123 TTCTGCGCGGCATAGACGTGTCTCTCCCAATGCCGATACAGATGTCTTTATTTAAC 1182
Qy      201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHisLeuYrAsnAlaGlyThr 220
Db      1183 ATCGACGATTCGAGACGATACGCAACGAACCCGAATGCCCTATACATGCCGAACA 1242
Qy      221 LeuYsAlaGlnThrIlyLeuGlyIlyIlePheAlaValAspArgThrAsnIlySlyLeuLeu 240
Db      1243 CTGAAGAGCCCAACAACTGGAATATTTCCAGTAGACAGAACCAATTAATAAATTTGCTC 1302
Qy      241 IleYsProIlyThrAsnAlaPheGlnAlaIleYrIlySgIlyuAsnYrAlaLeuThrMet 260
Db      1303 ATCAAAACCCAAACCAATGCTTTGAAAGCTCCCTAATAAGAAATTAAGCATTTGTGATG 1362
Qy      261 GilyProYrIlySValSerIleGlyIlyLeuYsProThrGlyIlyLeuMetValAspPheSer 280
Db      1363 GGGCGGCTTAAGTAAGCAAGCAATCAACCGACGAGAGATTTATGTCGATTTCTCC 1422
Qy      281 AspIleArgProYrGlyAsnHisThrGlyAsnSerAlaProSerValGlnAlaAspAsn 300
Db      1423 GATATCCGGCCATACGGCAATCATACGGGTAACTCCGCCCATCGTACAGGCTGATTAAC 1482
Qy      301 SerHisGlyIlyYrGlyYrSerAspGlnAlaValArgGlnHisArgGlnGlyIlyPro 320

```

Db	1483	AGTCATGAGGGGTATGATACAGCGATGAAGCACTGCGACACATAGACAAGGCAACCT	1542
RESULT 5	AF142582	AF142582	
LOCUS	AF142582	3294 bp	DNA
DEFINITION	Neisseria gonorrhoeae maf gene cluster, complete sequence.		
ACCESSION	AF142582		
VERSION	AF142582.1	GI:4838552	
KEYWORDS			
SOURCE	Neisseria gonorrhoeae		
ORGANISM	Neisseria gonorrhoeae		
REFERENCE	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria.		
AUTHORS	1 (bases 1 to 3294) Bickernjaeger,S., Meyer,T.F., Fischer,E., Maier,J., Manning,P.A., Rudel,T.G. Scheuerplug,I., Schulz,E. and Schwan,E.T.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-APR-1999) Molekulare Biologie, MPI fuer Infektionsbiologie, Monbijoustr. 2, Berlin 10117, Germany		
FEATURES	location/Qualifiers		
source	1..3294		
	/organism="Neisseria gonorrhoeae"		
	/mol_type="genomic DNA"		
	/strain="MS11"		
	/db_xref="taxon:485"		
gene	143..457		
	/gene="mafI"		
CDS	143..457		
	/gene="mafI"		
	/codon_start=1		
	/transl_table=11		
	/product="MafI"		
	/protein_id="ADJ1037.1"		
	/db_xref="GI:4838553"		
	/translation="MRSPRLRYKONKLIELFVAGVYARTAAEPDSIVYDDCYRRYDVLDGERSHFRINSHTHFAERKHINIGINFWAKRHLKRFDEIPEHEPYLKECHERRP"		
gene	590..1552		
	/gene="mafA"		
CDS	590..1552		
	/gene="mafA"		
	/codon_start=1		
	/transl_table=11		
	/product="adhesin MafA"		
	/protein_id="ADJ1038.1"		
	/db_xref="GI:4838554"		
	/translation="MRARLLIPLTSGVFLTASGTLTGIPSHGGRFAVEQELVAASARAAYKMDMDQALHGKVALYITMGDQSGSLTGGRYSIDALIREYINSPAVRDYTYPRVETTAFTSGGLGLTSLSTINAPALSTRQDSGSRSLGINTGMDYRNETLTNTPRDPAFLSHLVQTVFELRGIDVSPANDTVFINIDVFTLRNTEMLTYNAETLKAQTKLEYFAVDRTNKGLIKPKTNAFEAAVKENYALMGPYRVSKGIKPTGLIMVDSRDIRPYENHNGNSAPSVEDADNSHEGYSDBAVROHRCOP"		
gene	1592..3121		
	/gene="mafB"		
CDS	1592..3121		
	/gene="mafB"		
	/codon_start=1		
	/transl_table=11		
	/product="adhesin MafB"		
	/protein_id="ADJ1039.1"		
	/db_xref="GI:4838555"		
	/translation="MNLPIDQKFMMLFAAISLIQITPSHANGLDARLRDMDQAKHYEPGKRYHLFGNARGVKNRVCANVQTFDAIVAPLPIPTHERGFPBGIIIGETHFSGHGEVHPSPFNHDSKSTSDPSGGVGFYVQLRTSGEIHHPADGYDGPQGGYPEBQKADITYSVHKGSTKTKRINTVPOAPSPSDMLKENAGASGFLSRADAEAGLIWENDPKDMRARMDIRIGIVGAVNPPLTSGFGVGVAITPSAVSPVYAAARLTQGIHNGINSPEAQDLAASLIQDSAPAVKDGINSADQMADHNITATQATALAVEAAGTAKGVVLENPTKMDVKNKTYEKPEAPKPMQTVDEGAVGNKPKPSTQGHSTHSDNNTLPAPYVKPDTSLSPGTIIQDRIRWTKSKFPEKSLNGHFKAKGEFGDITIEDYQKASDLISKQSDKILIGQYTBHRRVRVDINNINIVLANPKTFIKITWFKPMIGREYVDGEFQKMG		

ORIGIN

**Alignment Scores:**

Pred. No.:	0	length:	3294
Score:	30.00	Matches:	320
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-617-835-4 (1-320) X AF142582 (1-3294

QY	1	MetaGAlaArgLeuLeuLeuLeuProLeuPheSerValPheIleuSerValCysGly	20
Db	590	ATGGGGGACGGCTGCTGATACCTATTCTTTTGGGTTTATTATTATCCGGCTCGGG	649
QY	21	ThrLeuThrglyIleProSerHisglYgIyGlyValArgPheAlaValgluIngleu	40
Db	650	ACACTGACAGGATATTCATCGCATGGCGGAGCAACGCTTCGGGGTCGACAGAACTT	709
QY	41	ValAlaAlaSerAlaArgAlaAlaValIysAspMetAspLeuGlnAlaIleuHisglYArg	60
Db	710	GTGGCGGCTTCGCGCAGAGCTGCGGTAAAGCATGGATTTCAGGCAATTACAGGACGA	769
QY	61	LysValAlaLeuTyrIleAlaIleThreGlyAspGlnGlySerGlySerLeuThrglyGly	80
Db	770	AAAGTGCATTTGTACATTGGCACTATGGCGCACCAGTTTCAGGCAATTTCAGGGGCT	829
QY	81	ArgTyrSerIleAspAlaIleuIleArgglYgIyTyrIleAsnSerProAlaValArgThr	100
Db	830	CGCTATCCATGATGACATGATTCGCGCGCAATACATTAACGCGCTTCGTCGGAC	889
QY	101	AspTyrThrTyrProArgTyrgluThrThrAlaGluThrThrSerGlyglYleuThrgly	120
Db	890	GATTACACCTATCCCGCGTTACGAAACACCCGCTGAACAACTACGCGGTTTGACGGGT	949
QY	121	LeuThrThrThrSerLeuSerThrIleuAsnAlaProAlaIleuSerArgThrGlnSerAspGly	140
Db	950	TTTAAACACTTCTTTATCTTACACTTAATGCGCTTCGACCTTCGCGCACCCCAATCAGAGCT	1000
QY	141	SerGlySerArgSerSerLeuGlyIleuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu	160
Db	1010	AGCGGAATAGGACAGCATGTGGGCTTAATAATTGGCGGAGTGGGGGATTATGAAATGA	1060
QY	161	ThrLeuThrThrAsnProArgAspThrAlaPheIleuSerHisIleuValGlnThrValPhe	180
Db	1070	ACCTTGACGACCAACCCGCGCAGACCTGCTTCTTCCCATTTGGTACAGACCGTATTT	1122
QY	181	PheIleuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleuAsn	200
Db	1130	TTCTCGCGCGGATAGAGCGTTGTTCTCTCGCAATGCCGATACAGATGTGTTATTAAAC	1189
QY	201	IleAspValPheGlyThrIleArgAsnArgThrGluMetHisIleuTyrAsnAlaGluThr	220
Db	1190	ATCGACGATTTTCGAAACGATACGCAACGAAACCGAATGCACCTTACATGCCGAACA	1244
QY	221	LeuIysAlaGlnThrIysLeuGluIuTyrPheAlaValAspArgThrAsnIysIleuLeu	240
Db	1250	CTGAAGCCCAACAAACTGGAAATATTTCGCAGTAGACAGAACCAATTAATAAATTGCTC	1309
QY	241	IleIysProIySerThrAsnAlaPheGluAlaAlaIuTyrIysGluAsnTyrAlaIeuThrPhe	260
Db	1310	ATCAAAACCCAAACCAATGCGCTTTAAGCTGCTTAAAGAAATTAACCATTTGGAGAG	1366
QY	261	GlyProTyrIysValSerIysglYIleIysProThrGluGlyIleuMetValAspPheSer	280
Db	1370	GGGCGGTTTAAGTAAGCAAGAGATCAAAACGACGAAAGATTAAATGGTGCATTTCTCC	1422
QY	281	AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGlnAlaAspAsn	300
Db	1430	GATATCCGGCCATACGGCAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATAAC	1488
QY	301	SerHisglYglYtyrGlytyrSerArgGlyAlaValArgGlnHisArgGlnIyGlnPro	320
Db	1490	AGTCATGAGGGGTATAGATACAGCATAGACGATGAGCGCAACATTAAGAAAGGGCAACT	1544

RESULT 6				
A96050	A96050	963 bp	DNA	linear
LOCUS				
DEFINITION	Sequence 83 from Patent WO924578.			
ACCESSION	A96050			
VERSION	A96050.1	GI:6779917		
PAT	07-SEP-2000			

## CONCLUSIONS

Alignment Scores:	
Pred. No.:	3,42e-290
Scores:	282.00
Percent Similarity:	99.38%
Best Local Similarity:	99.38%
Query Match:	86.12%
JB:	6
Length:	963
Matches:	319
Conservative:	0
Mismatches:	1
Indels:	2
Gaps:	0

US-10-617-835-4 (1-320) x A96050 (1-963)

Oy	1	MetArGlaaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerIaCyGly	20
Db	1	ATGGGGGACGCGCTGATACCTAATCTTTTATGATTTTAAATTTATACCCGCTGGGG	60
Oy	21	ThrLeuThrGlyIleProSerHisGlyGlyGlyLysArgPheAlaValGluGlnIleu	40
Db	61	ACACTGACGAGTATTTCCATCCGATGCGGAGGCAAAACGCTTCCGGTCSAACAGAACTT	120
Oy	41	ValAlaIaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg	60
Db	121	GTGGCGCGCTTCCGCAGAGCTGCGGTTAAAGCATGATTTTCACGACATTACACGAGCA	180
Oy	61	LysValAlaLeuIleuIleAlaThrMetCysAspGlnGlySerGlySerLeuThrGlyGly	80
Db	181	AAAGTGTGATTGTACATTGCAATTCAGCAATAGGGCGACCAAGGTTCCGGCAGTTTGACAGGGG	240
Oy	81	ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr	100
Db	241	CGTACTTCATTGATGACCTGATTTGCGGCGAATCATTAACAGCCCTCCGCTCCGACC	300
Oy	101	AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyLeuLeuThrGly	120
Db	301	GATTACACCTATCCGCGTTACGAAACCAACCGCTGAACCAATCAGCGGTTTGACGGGT	360
Oy	121	LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly	140
Db	361	TTAACACATCTTTATCTTACACTTAATCCCTCGACCTCGCGGACCCCAATCAGCGGT	420
Oy	141	SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu	160
Db	421	AGCGGAAGTACGACACGCTGGGCTTAATATTGGCGGGATGGGGGATTATCGAAATGA	480
Oy	161	ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe	180
Db	481	ACCTTGACGACCAACCCGCGGACACGCTCTTCTTCCCACTTGTCGACGACCGATATT	540
Oy	181	PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn	200
Db	541	TTCTCGCGCGGCATAGACGTTGTTCTCTCTGCAATGCCGATACAGATGTTTATTAAAC	600

Qy 201 ILeaspValPheglYThrIleArgAsnArgThrgluMetHisLeuTYrAsnAlaGluThr 220  
Db 601 ATCGACGATTTCGGAACATACGCAACGAAACCGAAATGACCTTATCAATGCCAAACA 660  
Qy 221 LeuylsAlaGlnThrLysLeuGluTYrPheAlaValAspArgThrAsnLysLysLeu 240  
Db 661 CTGAAGGCCCAACAAACCTGGAATATTTCGCAGTAGACAGAACCAATAAAAAATTGCTC 720  
Qy 241 IleYsProYsThrAsnAlaPheGluAlaAlaTYrLysGluAsnTYrAlaLeuTYrMet 260  
Db 721 ATCAAAACCCAAACCAATGCGTTGAACTGCTTAAAGAAATTTAGCATTTGGAGTAG 780  
Qy 261 GLYProTYrLysValSerLysGlyIleYsProThrgluGlyLeuMetValAspPheSer 280  
Db 781 GGGCGGCTTAAAGTAGCAAGAAAGATCAACCGACGGAAGATTGATGTGATTCTCC 840  
Qy 281 AspIleArg-ProTYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAs 300  
Db 841 GATATCCA-ACCATACGGCAATCATACGGGTAACTCCGCCCATCGTAGAGGCTGATAA 899  
Qy 300 nSerHisGluGlyTYrGlyTYrSerAspGluAlaValArgGlnHisArgGlnGlyInPr 320  
Db 900 CAGTCATGAGGGGTATGATACAGCATGAAGCAGTGCAACATAGACAAAGGCAACC 959  
Qy 320 o 320  
Db 960 T 960

RESULT 7  
AX043997 963 bp DNA linear PAT 24-NOV-2000  
LOCUS Sequence 76 from Patent WO0066791.  
DEFINITION AX043997  
ACCESSION AX043997.1 GI:11342888  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Neisseria gonorrhoeae  
Neisseria gonorrhoeae  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.

REFERENCE  
AUTHORS  
1 Piazza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,  
Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarlatti,M.,  
Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.  
Neisseria genomic sequences and methods of their use  
Patent: WO 0066791-A 76 09-NOV-2000;  
JOURNAL CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)  
FEATURES  
source  
1. .963  
Location/Qualifiers  
/organism="Neisseria gonorrhoeae"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:485"

ORIGIN  
Alignment Scores:  
Pred. No.: 3,42e-290 Length: 963  
Score: 282.00 Matches: 319  
Percent Similarity: 99.38% Conservative: 0  
Best Local Similarity: 99.38% Mismatches: 1  
Query Match: 88.12% Indels: 2  
DB: Gaps: 0

US-10-617-835-4 (1-320) x AX043997 (1-963)

Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
Db 1 ATGGCGGACGCGCTGCTGATACCTATTCCTTTTTCAGTTTAAATTTCGCCGCCGCGGG 60  
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGlnGluLeu 40  
Db 61 ACACCTGACAGGATATTCATTCGCATAGCGGAGGCAAAACCTTCGCGGTGAAACAAGACTT 120  
Qy 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60

Db 121 GTGGCCGCTTCGCAAGAGCTGCCGTTAAACAATGATTTACAGGCACTTACAGGACGA 180  
Qy 61 LysValAlaLeuTYrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
Db 181 AAAGTTGATTTGATTCATTGCAACTATGAGGCCACCAAGGTTAGCAGATTACAGGGGGT 240  
Qy 81 ArgTYrSerIleAspAlaLeuIleArgGlyGluTYrIleAsnSerProAlaValArgThr 100  
Db 241 CGCTACTCATTTGATTCATTCATTCGCGGCAATACATTAACAGCCCTCGCGTCGCAAC 300  
Qy 101 AspTYrThrTYrProArgTYrGluThrThrAlaGluThrThrSerGlyLysLeuThrGly 120  
Db 301 GATTACACCTATCCGCGTTACGAACCAACCCGCTGAAACAAATCATGAGGCGTTTACCGG 360  
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrglnSerAspGly 140  
Db 361 TTAAACCACTTCTTATCTTACACTTAATGCCCTTCACCTCTGCGCACCAATCAGACGGT 420  
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTYrArgAsnGlu 160  
Db 421 AGCGGAAGTAGAGACAGCTGGGCTTAAATTTGGCGGAGTGGGGATTATCGAAATGA 480  
Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
Db 481 ACCTTGACGACCAACCGCGCGACACTGCTTCTTCCCACTTGTCAGACCGGTATTT 540  
Qy 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
Db 541 TTCCTGGCGGCAATGACGTTGTTCTCTGCCAATGCCATACAGATGTTTATTAATAC 600  
Qy 201 ILeaspValPheglYThrIleArgAsnArgThrgluMetHisLeuTYrAsnAlaGluThr 220  
Db 601 ATCGACGATTTCGGAACGATACGCAACGAAACCGAAATGACCTTATCAATGCCAAACA 660  
Qy 221 LeuylsAlaGlnThrLysLeuGluTYrPheAlaValAspArgThrAsnLysLysLeu 240  
Db 661 CTGAAGGCCCAACAAACCTGGAATATTTCGCAGTAGACAGAACCAATAAAAAATTGCTC 720  
Qy 241 IleYsProYsThrAsnAlaPheGluAlaAlaTYrLysGluAsnTYrAlaLeuTYrMet 260  
Db 721 ATCAAAACCCAAACCAATGCGTTGAAAGCTGCTTAAAGAAATTTACGCAATTGTGATG 780  
Qy 261 GLYProTYrLysValSerLysGlyIleYsProThrgluGlyLeuMetValAspPheSer 280  
Db 781 GGGCGGCTTAAAGTAGCAAGAAAGATCAACCGACGGAAGATTGATGTGATTCTCC 840  
Qy 281 AspIleArg-ProTYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAs 300  
Db 841 GATATCCA-ACCATACGGCAATCATACGGGTAACTCCGCCCATCGTAGAGGCTGATAA 899  
Qy 300 nSerHisGluGlyTYrGlyTYrSerAspGluAlaValArgGlnHisArgGlnGlyInPr 320  
Db 900 CAGTCATGAGGGGTATGATACAGCATGAAGCAGTGCAACATAGACAAAGGCAACC 959  
Qy 320 o 320  
Db 960 T 960

RESULT 8  
A96046 963 bp DNA linear PAT 07-SEP-2000  
LOCUS Sequence 79 from Patent WO9924578.  
DEFINITION A96046  
ACCESSION A96046.1 GI:6779915  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified  
unidentified  
unclassified.

REFERENCE  
AUTHORS  
1 Piazza,M., Scarlato,V., Rappuoli,R., Grandi,G. and Masignani,V.  
TITLE Neisserial antigens  
Patent: WO 9924578-A 79 20-MAY-1999;  
JOURNAL



PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);  
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)

FEATURES  
Source

Location/Qualifiers  
1..963  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.:	2,99e-210	Length:	963
Score:	207.00	Matches:	307
Percent Similarity:	99.35%	Conservative:	0
Best Local Similarity:	99.35%	Mismatches:	1
Query Match:	64.69%	Indels:	2
DB:	6	Gaps:	0

US-10-617-835-4 (1-320) x A96046 (1-963)

```

OY 3 A1AATGLeuLeu11ePro11eLeuPheSerVal1Phe11eLeuSer1aCy6g1yThrLeu 22
Db 7 GCACGGCTGCTGATACCTAATCTTTTTCAGTTTATTTTATTCGCCCTGCGGACACTG 66
OY 23 ThrG1y1eProSerH1eG1yG1y1ySaRgPheAlaVal1G1uG1nG1uLeuValAla 42
Db 67 ACAGGTATTCCTCCGATGCGGAGGTAAACCTTTGCGGTGCAACAGAACTTGTGCC 126
OY 43 A1Aser1aArgAla1aVal1ySaPhe1aPheLeuG1nAlaLeuH1eG1yArG1ySVal 62
Db 127 GCTTCTGCAGAGCTGCGCTTAAGACATGATTTACAGGCATTTACACGAGCAAAAGTT 186
OY 63 A1aLeu1yTr11eAlaThMeG1yVaPnG1nG1ySeRg1ySe1euthrG1yG1yArG1y 82
Db 187 GCATTGTACATTGACCATATGCGGACCAAGGTTGACAGGATTTGACAGGCGGTGCTAC 246
OY 83 Ser11eAPa1aLeu11eArg1yG1u1yTr11eAenSerProAlaValArG1yThraP1y 102
Db 247 TCCATTGATGACCTATTCGTGCGAATACATTAACAGCCCTGCGCTGACCATTC 306
OY 103 Thr1yProArG1yG1uThrThra1aG1uThrThraSerG1yG1yLeuThrG1yLeuThr 122
Db 307 ACCTATCCACGTTACGAACCAACCGCTGAAACAATCATCGGCGGTGACAGTTTAAACC 366
OY 123 ThrSer1eSerTh1eLeuAenAlaProAlaLeuSer1aG1nSer1aPnG1ySeRg1y 142
Db 367 ACTTTTATATCACTTAATGCCCCCTGACCTCTCCGACCAATCAACACGATGACGA 426
OY 143 SerArG1ySer1eLeuG1yLeuAen11eG1yG1yMeG1yAaP1yArG1yAenG1yThr1e 162
Db 427 AGTAA-AGCACTGCGGCTTAAATTTTGCGGAGTGGGGATTTATCGAAATGAAACTT 485
OY 162 uThrThraPnProArG1yThra1aPheLeuSerH1eS1eUVal1G1nThrVal1PhePhe1e 182
Db 486 GACGACTAACCCGCGACACTGCTTTCTTCCACTTGATGACAGACCGTATTTTCTT 545
OY 182 uArG1y11eAPVal1a1SerProAlaAenAlaPnThraPVal1Phe11eAen11eAS 202
Db 546 GCGCGCATAGACGTTGTTCTCCCTGCAATGCCATACAGATGTTATTATTAACATCA 605
OY 202 PVal1PheG1yThr11eArG1yAenArG1yG1uMeH1eS1eU1yArAenAlaG1uThrLeu1y 222
Db 606 CGTATTCGGAACGATACGCAACAGAACCAATGACCTATCAATCCGGAACACTGA 665
OY 222 BA1aG1nThr1yS1eU1yTr1PheAlaVal1aPnArG1yThraen1yS1yLeuLeu11e1y 242
Db 666 AGCCCAACCAAACTGCAATATTTCAGATAGACAGAACCAATTAATAAATTCTCATCA 725
OY 242 aPro1yThraPnAlaPheG1uAlaAla1yTr1yS1eU1yAen1yTrAla1euthrMeG1yPr 262
Db 726 ACCAAACCAATGCGTTGAGCTGCTATTAAGAAATTAACGATTTGATGGGGCC 785
OY 262 cTyTr1ySa1Ser1yG1y11eYpProThrG1uG1yLeuMeCVal1aPnPhaSer1aP11 282

```

```

Db 786 GTATAAGTAACCAAGAAATTAACCGACGAGATTAATGTCGATTCCTCGATAT 845
OY 282 eArProTyR1yAenH1eTh1yAenSer1aProSerVal1G1uAlaAaPnSerH1 302
Db 846 CCGACCAATACCGCAATCATATACCGGTAACTCCGCCCATCCGTAGAGGCTGATTAACGTCA 905
OY 302 eG1uG1yTr1yTr1ySer1aPnG1u 310
Db 906 TGAAGGATGATGATACAGCATGAA 930

```

RESULT 9  
AX043995 963 bp DNA linear PART 24-NOV-2000  
LOCUS  
DEFINITION  
Sequence 74 from Patent WO0066791.  
ACCESSION  
AX043995  
VERSION  
AX043995.1 GI:11342887  
KEYWORDS  
SOURCE  
ORGANISM

Neisseria meningitidis  
Neisseria meningitidis  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.

REFERENCE

AUTHORS

Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C.,  
Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarlato, V.,  
Scarlato, V., Rappuoli, R., Frazer, C.M. and Grandi, G.  
Neisseria genomic sequences and methods of their use  
Patent: WO 006791-A 74 09-NOV-2000;  
JOURNAL  
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES

Source

1..963  
/organism="Neisseria meningitidis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:487"

ORIGIN

Alignment Scores:

Pred. No.:	2,99e-210	Length:	963
Score:	207.00 <td>Matches:</td> <td>307 </td>	Matches:	307
Percent Similarity:	99.35% <td>Conservative:</td> <td>0 </td>	Conservative:	0
Best Local Similarity:	99.35% <td>Mismatches:</td> <td>1 </td>	Mismatches:	1
Query Match:	64.69% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	6 <td>Gaps:</td> <td>0 </td>	Gaps:	0

US-10-617-835-4 (1-320) x AX043995 (1-963)

```

OY 3 A1AATGLeuLeu11ePro11eLeuPheSerVal1Phe11eLeuSer1aCy6g1yThrLeu 22
Db 7 GCACGGCTGCTGATACCTAATCTTTTTCAGTTTATTTTATTCGCCCTGCGGACACTG 66
OY 23 ThrG1y1eProSerH1eG1yG1y1ySaRgPheAlaVal1G1uG1nG1uLeuValAla 42
Db 67 ACAGGTATTCCTCCGATGCGGAGGTAAACCTTTGCGGTGCAACAGAACTTGTGCC 126
OY 43 A1Aser1aArgAla1aVal1ySaPhe1aPheLeuG1nAlaLeuH1eG1yArG1ySVal 62
Db 127 GCTTCTGCAGAGCTGCGCTTAAGACATGATTTACAGGCATTTACACGAGCAAAAGTT 186
OY 162 uThrThraPnProArG1yThra1aPheLeuSerH1eS1eUVal1G1nThrVal1PhePhe1e 182
Db 486 GACGACTAACCCGCGACACTGCTTTCTTCCACTTGATGACAGACCGTATTTTCTT 545
OY 182 uArG1y11eAPVal1a1SerProAlaAenAlaPnThraPVal1Phe11eAen11eAS 202
Db 546 GCGCGCATAGACGTTGTTCTCCCTGCAATGCCATACAGATGTTATTATTAACATCA 605
OY 202 PVal1PheG1yThr11eArG1yAenArG1yG1uMeH1eS1eU1yArAenAlaG1uThrLeu1y 222
Db 606 CGTATTCGGAACGATACGCAACAGAACCAATGACCTATCAATCCGGAACACTGA 665
OY 222 BA1aG1nThr1yS1eU1yTr1PheAlaVal1aPnArG1yThraen1yS1yLeuLeu11e1y 242
Db 666 AGCCCAACCAAACTGCAATATTTCAGATAGACAGAACCAATTAATAAATTCTCATCA 725
OY 242 aPro1yThraPnAlaPheG1uAlaAla1yTr1yS1eU1yAen1yTrAla1euthrMeG1yPr 262
Db 726 ACCAAACCAATGCGTTGAGCTGCTATTAAGAAATTAACGATTTGATGGGGCC 785
OY 262 cTyTr1ySa1Ser1yG1y11eYpProThrG1uG1yLeuMeCVal1aPnPhaSer1aP11 282

```

Qy 143 SerArg-SerSerLeuGlyLeuAenIleGlyIyMeGlyAspTyrArgAenIuThIle 162  
Db 427 AGTAA-ANGAGCTCGGCTTAATATTGGCGGATGGGGAATTTCGAAATGAACCTT 485  
Qy 162 uThrThsAsnProArgrAspThraIaPheLeuSerHisLeuValGlnThraValPhePheLe 182  
Db 486 GAGACATNAACCCGGCGGACATCGCTTCTTCTCCACTTGGTACGAGACGATATTTTCT 545  
Qy 182 uArgGlyIleAspValIserProAlaAsnAlaAspThrAspValPheIleAsnIleAs 202  
Db 546 GCGGGGATAGACGTTGTTCTCGCAATGCCATACAGATGTTATTAAATCAATCGA 605  
Qy 202 pValPheGlyThrIleArgAsnArgThnGluMeHisLeuTyrAsnAlaGluThrLeuLy 222  
Db 606 CGTATTCGGAACGATACCGACAGAACCGAAATGACCTTAACAAATGCCGAACACTGA 665  
Qy 222 sAlaGlnThrIleuGlyLeuTyrPheAlaValAspArgThraAsnIlySlyLeuLeuIleLy 242  
Db 666 AGCCCAACAAACCTGGAATATTTCGACGTAGACAGAACCAATAAAAATGCTCATCAA 725  
Qy 242 sProLyThrAsnAlaPheGluAlaIaTyrLyGluAsnTyrAlaLeuTyrMetGlyPr 262  
Db 726 ACCAAAAACCAATCGTTTGAAGCTGCTATTAAGAAATTAACGCAATGATGATGGGGCC 785  
Qy 262 oTyrLyValSerIySlyLeuSlyProThGluGlyLeuMetValAspPheSerAspI 282  
Db 786 GTATTAAGTAAAGCAAGAAATTAACCGAAGAGATTAATGTCATTTCTCCGATAT 845  
Qy 282 eArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerHi 302  
Db 846 CCGACCATACCGGCATCATACGGGTAACTCCGCCCATCGTAGAGCTGATAACAGTCA 905  
Qy 302 sGluGlyTyrGlyTyrSerAspGlu 310  
Db 906 TGAGGGGTATGATACAGCGATGAA 930

RESULT 10  
AE002559 10869 bp DNA linear BCT 25-MAY-2000  
LOCUS Neisseria meningitidis serogroup B strain MC58 section 201 of 206  
DEFINITION  
ACCESSION AE002559 GI:7413482  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Neisseria meningitidis MC58  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.  
REFERENCE  
AUTHORS  
1 (bases 1 to 10869)  
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,  
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,  
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,  
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blait, E.,  
Cittone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H.,  
Qin, H., Vamathevan, J., Gill, J., Scariato, V., Masignani, V.,  
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,  
Rappuoli, R. and Venter, J.C.  
Complete genome sequence of Neisseria meningitidis serogroup B  
strain MC58  
Science 287 (5459), 1809-1815 (2000)

JOURNAL  
MEDLINE  
PUBMED  
10710307  
AUTHORS  
2 (bases 1 to 10869)  
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,  
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,  
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,  
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,  
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blait, E.,  
Cittone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouri, H.,  
Qin, H., Vamathevan, J., Gill, J., Scariato, V., Masignani, V.,  
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,  
Rappuoli, R. and Venter, J.C.

TITLE  
JOURNML  
COMMENT  
FEATURES  
source  
Direct Submission  
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
On Apr 4, 2000 this sequence version replaced gi:7227357.  
Location/Qualifiers  
1..10869  
/organism="Neisseria meningitidis MC58"  
/mol\_type="genomic DNA"  
/strain="MC58"  
/db\_xref="taxon:122586"  
/note="serogroup: B"  
150..524  
/gene="NM2095"  
150..524  
/gene="NM2095"  
/note="similar to PID:1817528 percent identity: 53.99;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=1  
/product="adhesin complex protein, putative"  
/protein\_id="AA042412.1"  
/db\_xref="GI:7227358"  
/translation="WKLTTHAISSAIALSSMAAAAGTDPYAKKTVSYVCOGKKY  
KVTVGFNRKQGLTYYASAVINGKRVQMPVNLDKSDNVEFTFGKGGVVGTVMDKSY  
RKQPIMTTADNQIVFRDCSPR"  
complement (712..2178)  
/gene="NM2096"  
complement (712..2178)  
/gene="NM2096"  
/note="similar to SP:P33940 PID:424051 GB:U00096 percent  
identity: 69.73; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=1  
/product="malate:quinone oxidoreductase"  
/protein\_id="AA042413.1"  
/db\_xref="GI:7227359"  
/translation="MAEATDVVLVGGGIMSATIGVLLKEDEPSWEITLIERLEDALE  
SSNANNAAGTSHALCELANYPAGANGIIPARALNIAEFPYRSQPMATLVAGKLE  
DNSFIPVPMHSVMNEDHCSYLOKRDARTOKLFFNMFSTDRNTISMAPIAGWK  
RDENQPIAAYSAEGTDVDRGLTROWVKYLOGKGVTEFNHVEDIKRSDGAWWK  
TADRNDGQLTTRFTLFGAGGALTTLOKSIIPGKGYCGRPVGLPPRNSPPT  
AEQHNKAVYQASVAGPMSVPHLDITNVGKRLMGFPYAGFRSNLKKQSLMDPL  
SIHNDNLVPMICAGMANMPLTKYLIGELRTKFERFASLLEYPEANPDWELLTTAQ  
RVQILKKDSEKGVLOFTEIIVAHADSLAALGASGASTAVPLMIRLMHOCPEPRA  
PSWEDRLKELVPGYIGIKENPERADRIIAYTAKVLDI"  
complement (2230..2319)  
/gene="NM2097"  
complement (2230..2319)  
/gene="NM2097"  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="AA042414.1"  
/db\_xref="GI:7227360"  
/translation="MLATHSNPETAAYSDTHHCPSLLRRRL"  
complement (2715..3305)  
/gene="NM2098"  
complement (2715..3305)  
/gene="NM2098"  
/note="conserved hypothetical protein; identified by  
Glimmer2; putative"  
/codon\_start=1  
/transl\_table=1  
/product="conserved hypothetical protein"  
/protein\_id="AA042415.1"  
/db\_xref="GI:7227361"  
/translation="MSEKILPVLSGLGVRLERPLDVHNETGLREAVCDGFWKLGVS  
APHDRYADYIGTALATRLAPAVVDEAGRGVTTAYHFEPOIPLDIDGTWYAAAS  
RRIRINCKIMILDSAFVLAFCVGRWPDIDINTASQRIERLGAEKDGLVLRHMLR  
KDSVSDTVVYSMLREDWCKRRLITGRLAGYGVQ"

gene complement(3366. .3947)  
/gene="NMB2099"  
CDS complement(3366. .3947)  
/gene="NMB2099"  
/note="conserved hypothetical protein; identified by  
Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF42416.1"  
/db\_xref="GI:7227362"  
/translation="MRNEKRALRRELGRSOMGRDYAAATVKINHLKRYIKGR  
KIGVYPMGKELDGFVRAQKGAELIYPIEPRSRMPTPYADGVQKERRGR  
AKHYPPARGRKKRPHNDLILPVYMGDRIGYRGQAGGYDALSLMKRYRLQKTV  
GVGFACQLVRLPEAHDRSLDGFVSEAGILCF"  
gene 4025. .4258  
/gene="NMB2100"  
4025. .4258  
CDS /note="NMB2100"  
/note="hypothetical protein; identified by Glimmer2;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAF42417.1"  
/db\_xref="GI:7227363"  
/translation="MMHIVAIIGYLFVAVMYSAQPSIARALIIYVFAVLPVTFVFT  
ITVRRRLMRQEQAESEQKAROKOSGTP"  
gene 4431. .5159  
/gene="NMB2101"  
4431. .5159  
CDS /gene="NMB2101"  
/note="similar to SP.P34831 GB.X53651 PID:296030 percent  
identity: 78.02; identified by sequence similarity;  
putative"  
/codon\_start=1  
/transl\_table=11  
/product="30S ribosomal protein S2"  
/protein\_id="AAF42418.1"  
/db\_xref="GI:7227364"  
/translation="MSQITMRQMIKAGVHFGHTRFNNPKMAQYIFGARANKIHIVLE  
KILPMFOAEAVRRLVANKGTVLFPVGTQRQARDIIEEATPAGPFDVRYMLGMLT  
NYKTKOSIKRLKEKTALENASGSPKLEIMORVKEKLELGGIKMMKGLPDA  
IFVIDTGYOKGTLVEAKLGIPIVAVVDNTPSPDVKYIVRGNDSSAALIRLXCGGIA  
DAVEGRKQALQETVAAAOEAAAE"  
gene 5288. .6142  
/gene="NMB2102"  
5288. .6142  
CDS /gene="NMB2102"  
/note="similar to GB.D26562 SP.P02997 GB.D13334 GB.V00343  
PID:1208943 percent identity: 69.01; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="elongation factor TS (EF-TS)"  
/protein\_id="AAF42419.1"  
/db\_xref="GI:7227365"  
/translation="MAETAKVADLRAATGICMECKKALVEAGNFDKAEILIRIK  
SGAKAGLAGRTAGVLAVALNGVNGALVEVNCETDFVADKAGVEEAFNFAKTAAR  
KKPAVEELSELVEARAIILAKLGEMNSVRFQVIDANQALVAILHGLATEGLVE  
YKGSADVARKIGIMHIVAAKPCQSEAEVDAETVEKERHITYEQALASGADPDAIKVY  
EGRIKPLAEITLNGQAFVMPDQVTAQPSKENGTEVLSFVRKYKGDDIEKKAVDYAA  
EVAALAAKY"  
gene 6359. .7078  
/gene="NMB2103"  
6359. .7078  
CDS /gene="NMB2103"  
/note="similar to GB.D26562 SP.P29464 GB.D13334 GB.X78809  
PID:1208944 percent identity: 75.16; identified by  
sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="uridylylate kinase"

/protein\_id="AAF42420.1"  
/db\_xref="GI:7227366"  
/translation="MTQIKYKRYLILKLSGSLMSDPGPNHDTIVTQGEIAEVK  
MGVQVIGVGGGNIIRGVSAQSGMDRRTAIYMGWATVMAALAKDLPETLIGIARV  
QSALSMQOIAETVAPKPAIQYLEBKVVI FPAAGSNPFPTDTAAALGAENNCVNL  
KATVNDGVYADPKPDPSATRYETITFEALLKILKVMADATFALCRRERKLNIVFGI  
AKEGSLKRVITGDEGLVHC"  
gene 7262. .8225  
/gene="NMB2104"  
/note="similar to GP:483854 percent identity: 99.38;  
identified by sequence similarity; putative"  
gene 8265. .9716  
/gene="NMB2105"  
8265. .9716  
CDS /gene="NMB2105"  
/note="similar to GP:483855 percent identity: 0.00;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
Alignment Scores:  
Pred. No.: 2,1e-195 Length: 10869  
Score: 194.00 Matches: 307  
Percent Similarity: 99.03% Conservative: 0  
Best Local Similarity: 99.03% Mismatches: 1  
Query Match: 60.62% Indels: 3  
DB: 1 Gaps: 0  
US-10-617-835-4 (1-320) x AE002559 (1-10869)  
QY 3 AAlaArgLeuLeuIleProIleuPheSerValPheIleLeuSerAlaCysGlyThrLeu 22  
DB 7268 GCAAGCGCTGCATACCTATCTTTTCAGTTTATTTATCCGCGTGGGACACTG 7327  
QY 23 ThrGlyIleProSerHisGlyGlyGlyIlySarPheAlaValGluGlnGluLeuValAla 42  
DB 7328 ACAGGATATTCATCGATCGAGCGAGGTAAACGCTTGGGTGGAACAAGATCTGTGCC 7387  
QY 43 AAlaSerAlaAlaGlyAlaAlaValIlySerMetSerLeuGlnAlaLeuHisGlyValArgIlyVal 62  
DB 7388 GCTTGTGCAAGCGCGCGGTAAAGACATGATTAAACAGCATTCACGAGCAAGAAAGTT 7447  
QY 63 AAlaLeuTyrIleAlaThrMetGlyValSerGlnIlySerGlySerLeuThrGlyIlyValArgTyr 82  
DB 7448 GCATTGTATTCATTCACCATATGGGCAACCAAGCTTAGGCACTTACAGGGGGGTCGCTA 7507  
QY 82 rSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThrAspTyr 102  
DB 7508 CTCCTATGATGACATGATTCGTGGGAAATACATAACAGCCCTGCCGTACCGATTA 7567  
QY 102 rThrTyrProArgTyrGlyIlyThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThr 122  
DB 7568 CACCTATCCAGCTTACGAAACCAACCGCTGAAACAACATCAGCGGCTTGAACGTTTAAC 7627  
QY 122 rThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgTrpGlnSerAspGlySerGly 142  
DB 7628 CACTTCTTATCTACATTAATATGCCCCCTGCACCTCTGCAACCAATCAGACGATAGCGG 7687  
QY 142 ySerArg-SerSerLeuGlyLeuAsnIleGlyIlyMetGlyValAspTyrArgAsnGlyThrL 162  
DB 7688 AAGTAA-AAGCAGCTGTGGCTTAATATATGGCGGATGGGAGATATGCAATGAAACCT 7746  
QY 162 eUThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePheL 182  
DB 7747 TGACGACTAACCCGGCGACACTGCTTCTTCTCCACTGTGATACAGACCGATTTTTC 7806  
QY 182 eUArgGlyIleAspValIlySerProAlaAsnAlaAspThrAspValPheIleAsnIleA 202  
DB 7807 TGCGCGGATAGAGCTTTCTCTCTGCAATGCCGATACAGATGTTATTAATCATCG 7866  
QY 202 sPValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGlyThrLeuL 222  
DB 7867 ACGTATTCGGAACGATACGCAACGAATGACCTATTAACAATGCCGAACACTGA 7926

Qy 222 ysaIagInThrlYsleuGluYrPhealValaAspArgThrAsnYsleuIleu 242  
Db 7927 AAGCCAAACAAATGGAAATATTTTCGAGTAGACAGAACCAATATAAAATTTGCCATCA 7986  
Qy 242 ySPoLYsThrAsnAlaPheGluAlaAlATrYlYsGluAsnTrAlaLeuTrpMetGlyP 262  
Db 7987 AACCAAAACCAATCGTTTGAAGCTGCTATTAAGAAAATTAGCATTTGGATGGGGC 8046  
Qy 262 rOTrYlYsValSerlySGlyTlYelYsProThrGluGlyLeuMetValAspPheSerAspI 282  
Db 8047 CGTATAAAGTAGAACAAAGAAATTAAACCGACGGAAGATTATATGCTCATTTCTCCGATA 8106  
Qy 282 leaSPProTrYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerH 302  
Db 8107 TCCGACCATACGGCATATACGGGTAATCCGCCCATCCGTAGAGCTGATTAACAATC 8166  
Qy 302 isGluGlyTrYrGlyTrYrSerAspGlu 310  
Db 8167 ATGAGGGGTATGATACAGCGATGAA 8192

RESULT 11  
AX044035 AX044035 172325 bp DNA linear PAT 24-NOV-2000  
LOCUS Sequence 114 from Patent WO0066791.  
DEFINITION AX044035  
ACCESSION AX044035  
VERSION AX044035.1 GI:11342919  
KEYWORDS  
SOURCE  
ORGANISM  
Neisseria meningitidis  
Neisseria meningitidis  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.  
REFERENCE  
AUTHORS  
1 Pizsa, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,  
Masiagnani, V., Galeotti, C., Mora, M., Ratti, G., Scarlato, V.,  
Scarlato, V., Rappuoli, R., Frazer, C. M. and Grandi, G.  
TITLE  
Neisseria genomic sequences and methods of their use  
JOURNAL  
Patent: WO 0066791-A 114 09-NOV-2000;  
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)  
FEATURES  
SOURCE  
Location/Qualifiers  
1..172325  
/organism="Neisseria meningitidis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:487"  
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to  
349980 349980 bases-seq 108: 300001 to 649980 349980  
bases-seq 109: 600001 to 949980 349980 bases-seq 110:  
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980  
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq  
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to  
227325 172325 bases"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,83e-194 Length: 172325  
Score: 194.00 Matches: 307  
Percent Similarity: 99.03% Conservative: 0  
Best Local Similarity: 99.03% Mismatches: 1  
Query Match: 60.62% Indels: 3  
DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x AX044035 (1-172325)

Qy 3 AlaArgLeuLeuIleProIleLeuPheSerValPheIleuSerAlaCYsGlyThrIleu 22  
Db 121698 GCACGGCTCGATACCTATCTTTTTCAGTTTATTTATTCGCGCTGGGGACATCG 121757

Qy 23 ThrGlyIleProSerHisGlyGlyGlyYsAspArgPheAlaValGluGluIleuValAla 42  
Db 121758 ACAAGTATTCATCGCATGGGAGAGTAAACGCTTTGGCGTCGAACAAACTTGTGGCC 121817

Qy 43 AlaSerAlaArgAlaValAlaValYsAspMetAspLeuGluAlaLeuHisGlyArgYsVal 62  
Db 121817

Db 121818 GCTTCTGCCAGAGCTGCCGTTAAAGACATGATTTTACAGCATTAACAGACGAAAGTT 121877  
Qy 63 AlaLeuTrYrIleAlaTrpMetGlyAspGluGlySerGlySerLeuThr-GlyGlyArgTy 82  
Db 121878 GCATTGTACATTTGCACATATGAGGAGCCAGAGCTTCAGCACTTTGACAGGGGGGCGCTA 121937  
Qy 82 rSerIleAspAlaLeuIleArgGlyGlyTrYrIleAsnSerProAlaValArgThrAspTy 102  
Db 121938 CTCCTATGATCATCATGATTCGTGGGGAATACATAAAACAGCCCTGCCGTCGATACGATTA 121997  
Qy 102 rThrTrYrProArgTrYrGlyTrYrThrAlaGluTrThrSerGlyGlyLeuTrYrGlyLeuTh 122  
Db 121998 CACCTATTCACGTACGAAACACCGCTGAAACAAACATCAGCGGTTGACAGCTTTAAC 122057  
Qy 122 rThrSerIleuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142  
Db 122058 CACTTCTTATCTACCTTAATATGCCCCCTGCACCTCTTCGACCCCAATACAGCGTAGCGG 122117  
Qy 142 ySerArg-SerSerIleuGlyLeuAsnIleGlyGlyMetGlyAspTyArgAsnGluThrL 162  
Db 122118 AAGTAA-AAGCAGCTGGGCTTAATATTTGGCGGGATGGGGATTATCGAAATGAACCT 122176  
Qy 162 eUTHrThrAsnProArgAspThrAlaPheLeuSerHisleuValGluThrValPhePheL 182  
Db 122177 TGACGACTAACCCCGCGACACTGCCCTTCTTCCACTTGTAACAGACCGTATTTTCC 122236  
Qy 182 eUaArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsnIleA 202  
Db 122237 TGCGGGCATAGACAGTGTGTTCTCTCGCAATGCCGATACGATGTGTTTATTACATCG 122296  
Qy 202 sPValPheGlyThrIleArgAsnArgThrGluMetHisleuTrYrAsnAlaGluThrLeuL 222  
Db 122297 ACGTATTCGGAACGATACGACAGAACGAAATGCACTTATCAATGCCGAAACACTGA 122356  
Qy 222 ysaIagInThrlYsleuGluYrPhealValaAspArgThrAsnYsleuIleu 242  
Db 122357 AAGCCAAACAAACCTGAAATATTTCCGAGTAGACAGAACCAATATAAAATTTGCCATCA 122416  
Qy 242 ySPoLYsThrAsnAlaPheGluAlaAlATrYlYsGluAsnTrAlaLeuTrpMetGlyP 262  
Db 122417 AACCAAAACCAATCGTTTGAAGCTGCTATTAAGAAAATTAGCATTTGGATGGGGC 122476  
Qy 262 rOTrYlYsValSerlySGlyTlYelYsProThrGluGlyLeuMetValAspPheSerAspI 282  
Db 122477 CGTATAAAGTAGAACAAAGAAATTAAACCGACGGAAGATTATGCTCATTTCTCCGATA 122536  
Qy 282 leaSPProTrYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerH 302  
Db 122537 TCCGACCATACGGCATATACGGGTAATCCGCCCATCCGTAGAGCTGATTAACAATC 122596  
Qy 302 isGluGlyTrYrGlyTrYrSerAspGlu 310  
Db 122597 ATGAGGGGTATGATACAGCGATGAA 122622

RESULT 12  
A96048 A96048 963 bp DNA linear PAT 07-SEP-2000  
LOCUS Sequence 81 from Patent WO924578.  
DEFINITION A96048  
ACCESSION A96048  
VERSION A96048.1 GI:6779916  
KEYWORDS  
SOURCE  
unidentified  
unidentified  
ORGANISM  
unclassified.  
REFERENCE  
AUTHORS  
1 Pizsa, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masiagnani, V.  
TITLE  
Neisserial antigens  
JOURNAL  
Patent: WO 924578-A 81 20-MAY-1999;  
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);  
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)  
FEATURES  
SOURCE  
Location/Qualifiers  
1..963  
/organism="unidentified"

## ORIGIN

/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

Alignment Scores: 2,096-180 Length: 963  
Pred. No.: 179.00 Matches: 279  
Score: 99.29% Conservative: 0  
Percent Similarity: 99.29% Mismatches: 1  
Best Local Similarity: 99.29% Indels: 2  
Query Match: 55.94% Gaps: 0  
DB: 6

US-10-617-835-4 (1-320) x A96048 (1-963)

QY 3 AAlaArgLeuLeu1LeuPro1LeuPheSerVal1Phe1LeuSerAlaCysGlyThrLeu 22  
DB 7 GCACGGCTGCTGATACCTATCTTTTTCAGTTTATTTATTCGCCCTGGCGGACACTG 66  
QY 23 ThrGly1LeuProSerH1eGlyGly1GlySerArgPheAlaVal1GluGlnGluLeuVal1Ala 42  
DB 67 ACAGGTATTCATCCATGCGATGCGGAGGTAAACCTTTGCGGTGAACAGAACTTGTGGCC 126  
QY 43 AAlaSerAlaArgAlaAlaVal1LysAspMetAspLeuGlnAlaLeuH1eGlyArgLysVal 62  
DB 127 GCTTCTGCAGAGCTGCGCTTAAAGACATGATTTACAGGCACTTACACGACGAAAGTT 186  
QY 63 AAlaLeuTyr1Lea1aThMetGly1AspGlnGlySerGlySerLeuThrGly1GlyArgTyr 82  
DB 187 GCATTGTACATTGCAACTATGAGCGGACCAAGGTTGAGGAGTTTGCAGGGGTCGCTAC 246  
QY 83 Ser1LeaAspAlaLeu1LeaArgGly1Gly1GlyLeaSerProAlaVal1ArgThrAspTyr 102  
DB 247 TCCATTGATGACCTATTCGTGGCGAATACATTAACAGCCCTGCGGTCCGTACCGATTAC 306  
QY 103 ThrTyrProArgTyrGly1UthrThra1GluThrTrsSerGly1GlyLeuThrGly1LeuThr 122  
DB 307 ACCTATCCACGCTTACGAACCAACCGCTGAACACATCAGGCGGCTTGCAGGTTTAC 366  
QY 123 ThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142  
DB 367 ACTTCTTATATCACTTAATACCCCTGCACTCGCGCACCAATACAGACCGTAGCGGA 426  
QY 143 SerArg-SerSerLeuGlyLeuAsn1eGly1GlyMetGlyAspTyrArgAsnGly1Uthr 162  
DB 427 AGTAA-AAGCACTGCGGCTTAAATATGCGGAGTGGGATTAATCGAAATGAACCTT 485  
QY 182 UthrThraAspProArgAspThrAlaPheLeuSerH1eLeuVal1GlnThrVal1PhePhe 182  
DB 486 GACGACTAACCCGCGGACACTGCTTTTCCCACTTGGTACAGACCGTATTTTCTCT 545  
QY 182 UArgGly1LeaAspVal1ValSerProAlaAsnAlaAspThrAspVal1Phe1LeaAsn1Lea 202  
DB 546 GCGGCGCATAGACGTTGTTCTCCGCGCAATGCCGATACGATGCTTATTAACATCGA 605  
QY 202 PVal1PheGly1Uthr1LeaArgAsnArgThrGluMetH1eLeuTyrAsnAlaGlu1UthrLeu 222  
DB 606 CGTATTCGGAACGATACCGCAACAGAACCCGAAATGACCTATACATGCGGAACACTGA 665  
QY 222 AAlaGlnThrLysLeuGluTyrPheAlaVal1AspArgThrAsnLysLeuLeu1eGly 242  
DB 666 AGCCCAACCAAACTCGAATATTTGCGACTACAGAACCAATTAATAAATCTCTCATCA 725  
QY 242 AspLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTyrMetGly1Yr 262  
DB 726 ACCAAACCAATGCGTTGAAGCTGCTATTAAGAAATTAACGATGTGATGGGACC 785  
QY 282 CTrLysVal1SerLysGly1LeuProThrGly1GlyLeuMetVal1AspPheSerAsp1 282  
DB 786 GATATAAGTAGCAAAAGAAATTAACCGACAGAAAGATTAATGTCGATTTCTCCGATAT 845  
QY 282 e 282  
DB 846 C 846

## RESULT 13

AX043999

LOCUS AX043999

DEFINITION Sequence 78 from Patent WO066791.

ACCESSION AX043999

VERSION AX043999.1 GI:11342889

KEYWORDS

SOURCE

ORGANISM

Neisseria meningitidis

Neisseria meningitidis

Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

Neisseriaceae; Neisseria.

REFERENCE

AUTHORS

1 Pizarro, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C.,

Maestri, V., Galeotti, C., Mora, M., Ratti, G., Scarfelli, M.,

Scarlato, V., Rappunli, R., Fraxer, C.M. and Grandi, G.

Neisseria genomic sequences and methods of their use

Patent: WO 0066791-A 78 09-NOV-2000;

CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES

source

1. 963

/organism="Neisseria meningitidis"

/mol\_type="unassigned DNA"

/db\_xref="taxon:487"

ORIGIN

Alignment Scores: 2,096-180 Length: 963

Pred. No.: 179.00 Matches: 279

Score: 99.29% Conservative: 0

Best Local Similarity: 99.29% Mismatches: 1

Query Match: 55.94% Indels: 2

DB: 6 Gaps: 0

US-10-617-835-4 (1-320) x AX043999 (1-963)

QY 3 AAlaArgLeuLeu1LeuPro1LeuPheSerVal1Phe1LeuSerAlaCysGlyThrLeu 22  
DB 7 GCACGGCTGCTGATACCTATCTTTTTCAGTTTATTTATTCGCCCTGGCGGACACTG 66  
QY 23 ThrGly1LeuProSerH1eGlyGly1GlySerArgPheAlaVal1GluGlnGluLeuVal1Ala 42  
DB 67 ACAGGTATTCATCCATGCGATGCGGAGGTAAACGCTTGGGCTGAACAGAACTTGTGGCC 126  
QY 43 AAlaSerAlaArgAlaAlaVal1LysAspMetAspLeuGlnAlaLeuH1eGlyArgLysVal 62  
DB 127 GCTTCTGCAGAGCTGCGCTTAAAGACATGATTTACAGGCACTTACACGACGAAAGTT 186  
QY 63 AAlaLeuTyr1Lea1aThMetGly1AspGlnGlySerGlySerLeuThrGly1GlyArgTyr 82  
DB 187 GCATTGTACATTGCAACTATGAGCGGACCAAGGTTGAGGAGTTTGCAGGGGTCGCTAC 246  
QY 83 Ser1LeaAspAlaLeu1LeaArgGly1Gly1GlyLeaSerProAlaVal1ArgThrAspTyr 102  
DB 247 TCCATTGATGACCTATTCGTGGGATACATTAACAGCCCTGCGGTCCGATTCGATTAC 306  
QY 123 ThrTyrProArgTyrGly1UthrThra1GluThrTrsSerGly1GlyLeuThrGly1LeuThr 122  
DB 367 ACTTCTTATATCACTTAATACCCCTGCACTCGCGCACCAATACAGACCGTATTTTCTCT 545  
QY 182 UthrThraAspProArgAspThrAlaPheLeuSerH1eLeuVal1GlnThrVal1PhePhe 182  
DB 427 AGTAA-AAGCACTGCGGCTTAAATATGCGGAGTGGGATTAATCGAAATGAACCTT 485  
QY 242 AspLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTyrMetGly1Yr 262  
DB 726 ACCAAACCAATGCGTTGAAGCTGCTATTAAGAAATTAACGATGTGATGGGACC 785  
QY 282 CTrLysVal1SerLysGly1LeuProThrGly1GlyLeuMetVal1AspPheSerAsp1 282  
DB 786 GATATAAGTAGCAAAAGAAATTAACCGACAGAAAGATTAATGTCGATTTCTCCGATAT 845  
QY 282 e 282  
DB 846 C 846



putative mobilization protein (200 aa), fasta scores: E() :  
0.551% identity in 167 aa overlap, and PIC\_ECOLI cell  
filamentation protein PIC (200 aa), fasta scores: E() :  
0.093, 29.4% identity in 126 aa overlap"

misc\_feature

gene

CDS

misc\_feature

gene

CDS

RBS

misc\_feature

```
/product="hypochemical protein NMA0004"  
/protein_id="CAB83324.1"  
/db_xref="GI:7378782"  
/db_xref="GOA:Q9JQR9"  
/db_xref="UniProt/TREMBL:Q9JQR9"  
/translation="MPSNPISGKTKMSIDEGSLHNRAPLESQIDRIEVTGAGLQO  
IHRVFGGLYDPAQOIRSDNTSKGFRFANAMYLAKELVKIEQMERPEEELIAYVE  
MNIAPRLENGRSTRITLDVLKKNLKVVMQVSKTLYIQMERSPVNDLEURL  
LKDNLTDVDNREILFKGIEQSYTEGTEKG"  
/complement(3516..3525)  
/gene="NMA0004"  
/note="Core DNA uptake sequence: gccgcctgaa"  
/label=DUS  
/complement(3561..3866)  
/gene="nuoK"  
/complement(3561..3866)  
/gene="nuoK"  
/EC number="1.6.5.3"  
/note="NMA0005, nuoK, NADH dehydrogenase I chain K, len:  
101 aa; similar to many e.g. NUO_K_RHOCA NADH dehydrogenase  
I chain K (EC 1.6.5.3) (102 aa), fasta scores: E() :  
1.1e-24; 65.0% identity in 100 aa overlap. Contains Pfam  
match to entry PF00420 oxidored q2,  
NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"  
/codon_start=1  
/transl_table=11  
/product="NADH dehydrogenase I chain K"  
/protein_id="CAB83325.1"  
/db_xref="GI:7378783"  
/db_xref="GOA:Q9JQU9"  
/db_xref="UniProt/TREMBL:Q9JQU9"  
/translation="MITTTHYIVLGALLFGISAMGIFPMRKXVVLMSIBMLAVN  
FNFIASFQHLGDTAGQIFVFFVITVAABASALGIALIMVLVTRNRTINVAIDELKG"  
/complement(3564..3863)  
/gene="nuoK"  
/note="Pfam match to entry PF00420 oxidored q2,  
NADH-ubiquinone/plastoquinone oxidoreductase chain 4L,  
score 141.50, E-value 1.5e-38"  
/complement(3863..4534)  
/gene="nuoJ"  
/complement(3863..4534)  
/gene="nuoJ"  
/EC number="1.6.5.3"  
/note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len:  
223 aa; similar to many e.g. NUOJ_RHOCA NADH dehydrogenase  
I chain J (EC 1.6.5.3) (202 aa), fasta scores: E() :  
9.7e-21; 41.3% identity in 201 aa overlap. Contains Pfam  
match to entry PF00499 oxidored q3,  
NADH-ubiquinone/plastoquinone oxidoreductase chain 6"  
/codon_start=1  
/transl_table=11  
/product="NADH dehydrogenase I chain J"  
/protein_id="CAB83326.1"  
/db_xref="GI:7378784"  
/db_xref="GOA:Q9JX90"  
/db_xref="UniProt/TREMBL:Q9JX90"  
/translation="MTFQILIFYPYPAVILILGATITVTAKPVAALILVITFCVSAM  
LMLMQAEFLGLTVVTVGAVWVLFVFMMLNIDISEMRGFRHAPVAGVGTLL  
AAVALILVNPKTDLAAGFKMDIPADVNNIRDLGSRITYTDLLFELAAVLLLGWV  
AAIALVHRKTVNPKRMDPADQYKVRADQGRMLVMEAVKPKQTESAESSEVSDLPK  
EEGKA"  
/complement(3874..3879)  
/gene="nuoJ"  
/complement(4034..4525)  
/note="Pfam match to entry PF00499 oxidored q3,  
NADH-ubiquinone/plastoquinone oxidoreductase chain 6,
```

score 77.90, E-value 2e-19"  
/complement(4570..5277)  
/gene="NMA0007"  
/complement(4570..5277)  
/gene="NMA0007"  
/note="NMA0007, unknown, len: 235 aa"

Alignment Scores:  
Pred. No.: 5,24e-178 Length: 340806  
Score: 179.00 Matches: 279  
Percent Similarity: 99.29% Conservative: 0  
Best Local Similarity: 99.29% Mismatches: 1  
Query Match: 55.94% Indels: 2  
DB: 1 Gaps: 0

US-10-617-835-4 (1-320) x NMA122491 (1-340806)

```
QY 3 ALaArgLeuLeuLeuProIleLeuPheSerValPheIleLeuSerAlaCysGlyThrLeu 22  
DB 306925 GCACGGCTGCTGATACCTATCTTTTTCAGTTTATTTATCCGCTGGCGACACG 306866  
QY 23 ThrGlyIleProSerHisGlyGlyGlyLysArgPheAlaValGlnGlnLeuValAla 42  
DB 306865 ACAGGATTCATTCGATCGGCGGAGGTAAACGCTTGGCGTCAACAGAACTGTGGCC 306806  
QY 43 AlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyValArgVal 62  
DB 306805 GCTTCTGCAGAGCTCGCGTTAAAGACATGATTTACAGCATTTACACGACGAAAGTT 306746  
QY 63 AlaLeuTrpIleAlaThrMetGlyValAspGlnLysSerGlySerLeuThrGlyValArgTyr 82  
DB 306745 GCATTGTATTCATTCGACATATGGCGACCAAGTTTCAGGCAATTTACAGGGGGTGGTAC 306686  
QY 83 SerIleAspAlaLeuIleArgGlyGlyTrpIleAsnSerProAlaValArgThrAspTyr 102  
DB 306685 TCCATTGATGACATGATTCGTGGCGAATACATTAACACCCGCTCGTACCATTAAC 306626  
QY 103 ThrTrpProArgTyrGlyThrThrAlaLuhThrSerGlyGlyLeuThrGlyLeuThr 122  
DB 306625 ACCTATCCACGTTACGAACACCGCTTAACACATCAGGGGGTGTACAGGTTTAAACC 306566  
QY 123 ThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142  
DB 306565 ACTTCTTATATCACTTAATAGCCCTTCACCTCTGGCACCCATTCAGAGGTCGCGA 306506  
QY 143 SerArg-SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlnThrIle 162  
DB 306505 AGTAA-AAGCAGCTCGGCTTAATATATGCGGGATGCGGATTCGAAATGAACCTT 306447  
QY 162 UThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePhe 182  
DB 306446 GACGACTAACCCGCGCGACACTGCTTCTTCTCCACTTGGTACGACCGATTTTCTT 306387  
QY 182 uArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsnIleAs 202  
DB 306386 GCGGGCGATAGACGTTGTTTCTCGCAATGCGCATCGGATGTATTATTAATCATCGA 306327  
QY 202 PValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGlnThrLeuLys 222  
DB 306326 CGTATTCGGAACGAATACCAACAGAACCAATATGACCTTAATACATGCGGAACCTGAA 306267  
QY 222 AlaGlnThrHisLeuGlnIleTyrPheAlaValAspArgThrAsnIleLysLeuLeuIleLys 242  
DB 306266 AGCCCAACCAAACTCGAATATTTTCGAGTACAGAACCAATTAATAATTCCTATCAA 306207  
QY 242 sProLysThrAsnAlaPheGlnAlaValLysArgGluAsnTyrAlaLeuTrpMetGlyPr 262  
DB 306206 ACCAATAAACCAATGCGTTGAAGCTGCTTAATGAATAATTAACGATGTGGATGGAGCC 306147  
QY 262 cTyrIleValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSerAspIle 282  
DB 306146 GTATTAAGTTAAGCAAAAGAAATTAACCAAGAAAGATTAATGTCGATTTCTCGGATAT 306087
```

QY 282 e 282  
Db 306086 C 306086

RESULT 15  
A96044  
LOCUS A96044 640 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 77 from Patent WO924578.  
ACCESSION A96044  
VERSION A96044.1 GI:6779914  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

REFERENCE 1  
AUTHORS Pizsa, M., Scarlato, V., Rappuoli, R., Grandi, G. and Maignani, V.  
TITLE Neisserial antigens  
JOURNAL Patent: WO 9924578-A 77 20-MAY-1999;  
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);  
CHIRON SPA (IT); GRANDI GUIDO (IT); MALIGNANI VEGA (IT)  
location/qualifiers  
1. 640  
source /organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

FEATURES  
source

ORIGIN

Alignment Scores:  
Pred. No.: 2.51e-63 Length: 640  
Score: 69.00 Matches: 123  
Percent Similarity: 98.40% Conservative: 0  
Best Local Similarity: 98.40% Mismatches: 1  
Query Match: 21.56% Indels: 2  
DB: Gaps: 0

US-10-617-835-4 (1-320) x A96044 (1-640)

QY 90 G|YGLuTYrIleAsnSerProAlaValArgThrAspTYrThrTYrProArgTYrGluThr 109  
|||  
Db 269 GGGCAATACATTAACAGCCCTGCCGTCCGTACCGATTACACTTACAGTAACGAAACC 328  
|||  
QY 110 ThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuAsn 129  
|||  
Db 329 ACCCGTAAACAAACATCAGCGGTTTGACAGTTTAAACACTTCTTATCTACACTTAAT 388  
|||  
QY 130 AlaProAlaLeuSerArgThrGlnSerAspGlySerArg-SerSerLeuGlyLe 149  
|||  
Db 389 GCCCTGCACCTCTCGCACCCCAATCAGCGGTAGCGAAGTAA-AGCGAGTCTGGGCTT 447  
|||  
QY 149 uAsnIleGlyGlyMetGlyAspTYrArgAsnGluThrLeuThrThrAsnProArgAspTh 169  
|||  
Db 448 AAATATTGGCGGATGGGGATTATCGAAATGAACCTTGACGACTAACCCGCGGACAC 507  
|||  
QY 169 rAlaPheLeuSerHisLeuValGlnThrValPhePheLeuArgGlyIleAspValValSe 189  
|||  
Db 508 TGCCTTCTTCCACCTTGATGATACAGACCGATTTTCTCGCGCGCATAGACGTGTTTC 567  
|||  
QY 189 rProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleArgAs 209  
|||  
Db 568 TCCCGCAATGCCGATAGATGATTTATTAACTGACGATATTCGGAACGATACGCA 627  
|||  
QY 209 nArgThrGluMet 213  
|||  
Db 628 CAGAACCGAAATG 640  
|||

Search completed: August 19, 2005, 03:26:13  
Job time : 4036 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 18, 2005, 23:17:19 ; Search time 653 Seconds

(without alignments)  
3183.789 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 1645

Sequence: 1 MRARLLRILPLSVFLSLACG.....SHEGYGSDAIVRQRCOP 320

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlp  
-O=/cgn2\_1/USPTO.spool.p/US1061735/Runat.18082005.115609.8398/app.query.fasta\_1.519  
-DB=Published Applications NA -OPMT=fastcap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-IOOPCL=0 -IOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=blowsum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=spc -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US1061735\_@CGN 1.1.723\_@Runat.18082005.115609.8398  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1645	100.0	1136	9 US-09-043-302-6	Sequence 6, Appl1
2	1645	100.0	1136	22 US-10-617-835-6	Sequence 6, Appl1
3	1645	100.0	1136	9 US-09-043-302-1	Sequence 1, Appl1
4	1645	100.0	3287	22 US-10-617-835-1	Sequence 69, Appl1
5	1029	62.6	939	17 US-10-330-880-69	Sequence 1, Appl1
6	107.5	6.5	9025608	15 US-10-156-261-1	Sequence 10, Appl1
7	104.5	6.4	1163020	17 US-10-398-221-10	Sequence 2058, Ap
8	104.5	6.4	3011208	17 US-10-398-221-2058	Sequence 87557, A
9	103.5	6.3	2604	19 US-10-437-963-87357	Sequence 16487, A
10	101	6.1	1335	17 US-10-282-122A-16487	Sequence 12803, A
11	101	6.1	1795	18 US-10-425-114-12803	Sequence 56736, A
12	101	6.1	1946	18 US-10-424-589-56736	Sequence 11877, A
13	99.5	6.0	1154	18 US-10-424-589-11877	Sequence 19283, A
14	99	6.0	2622	17 US-10-282-122A-19283	Sequence 69, Appl1
15	98.5	6.0	939	17 US-10-146-772-69	Sequence 69, Appl1
16	98.5	6.0	939	17 US-10-241-742-69	Sequence 69, Appl1
17	98.5	6.0	939	17 US-10-440-523-69	Sequence 69, Appl1
18	98.5	6.0	939	18 US-10-440-503-69	Sequence 69, Appl1
19	98.5	6.0	939	18 US-10-461-925-69	Sequence 1, Appl1
20	98.5	6.0	90597	21 US-10-211-028-1	Sequence 6607, Ap
21	98	6.0	1665	16 US-10-032-585-6607	Sequence 1, Appl1
22	97	5.9	36602	21 US-10-762-107-1	Sequence 3576, Ap
23	96.5	5.9	2149	17 US-10-398-221-3576	Sequence 6195, Ap
24	96.5	5.9	2250	18 US-10-425-114-6195	Sequence 47132, A
25	96	5.8	1422	17 US-10-369-493-47132	Sequence 14789, A
26	95.5	5.8	1195	20 US-10-425-115-41689	Sequence 1, Appl1
27	95.5	5.8	2100	17 US-10-282-122A-14771	Sequence 51, Appl1
28	95.5	5.8	64492	18 US-10-378-083-1	Sequence 157, App
29	95	5.8	2058	18 US-10-335-977-51	Sequence 157, App
30	94.5	5.7	1089	9 US-09-071-035-157	Sequence 157, App
31	94.5	5.7	1089	14 US-10-206-576-157	Sequence 157, App
32	94.5	5.7	1089	21 US-10-912-352-157	Sequence 485, App
33	94.5	5.7	1929	9 US-09-070-927A-485	Sequence 8649, Ap
34	94.5	5.7	2025	18 US-10-425-114-6649	Sequence 110807, A
35	94.5	5.7	2312	18 US-10-424-589-110807	Sequence 17669, A
36	94	5.7	1293	17 US-10-282-122A-17669	Sequence 14952, A
37	94	5.7	1461	19 US-10-437-963-14952	Sequence 1, Appl1
38	94	5.7	9025608	15 US-10-156-761-1	Sequence 5949, Ap
39	93.5	5.7	3000	15 US-10-156-761-5949	Sequence 14432, A
40	93.5	5.7	3222	17 US-10-282-122A-14432	Sequence 2, Appl1
41	93.5	5.7	3816	21 US-10-875-518-2	Sequence 12158, A
42	93	5.7	1038	19 US-10-437-963-12158	Sequence 15413, A
43	93	5.7	1356	17 US-10-282-122A-15413	Sequence 13676, A
44	93	5.7	1404	19 US-10-767-701-13676	Sequence 30610, A
45	93	5.7	2253	17 US-10-282-122A-30610	

#### ALIGNMENTS

RESULT 1  
US-09-043-302-6  
; Sequence 6, Application US/09043302  
; Publication No. US20020086349A1  
; GENERAL INFORMATION:  
; APPLICANT: MEYER, Thomas F.  
; APPLICANT: RUDEL, Thomas  
; APPLICANT: SCHUEERPLUG, Ina  
; APPLICANT: MAIER, Jurgen  
; APPLICANT: EICKERNJAGER, Sandra  
; APPLICANT: SCHMIDT, Thomas  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
; FILE REFERENCE: 0147-172P  
; CURRENT APPLICATION NUMBER: US/09/043,302  
; CURRENT FILING DATE: 1998-06-08

```
; EARLIER APPLICATION NUMBER: PCT/EP96/04092
; EARLIER FILING DATE: 1995-09-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1094)
US-09-043-302-6

Alignment Scores:
Pred. No.: 3,21e-193 Length: 1136
Score: 1645.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-617-835-4 (1-320) x US-09-043-302-6 (1-1136)

Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db 135 ATGCGGCGACGGCTGCTGATACCTATCTTTTTCGTTTATTTATTCGCCCTGCGGG 194
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGlnGlnLeu 40
Db 195 ACACGACAGGATATTCATCGCATGCGAGGCAAACTTCGCGGTGACAACTT 254
Qy 41 ValAlaIaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60
Db 195 ATGCGCGCTTCTGCGACAGCTCCGCTTAAAGACATGATTTACAGCATTTACAGGACGA 314
Qy 255 GTGGCGCTTCTGCGACAGCTCCGCTTAAAGACATGATTTACAGCATTTACAGGACGA 314
Db 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Db 315 AAAGTTGATTTGATTCATTCGCAACTATGGGCGACCAAGTTTCAGGACGTTTGACAGGGGT 374
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
Db 375 CGCTACCTCATTTGATGACCTGATTCGCGCGCAATACATTAACAGCCGCTCGGACCC 434
Qy 101 AspTyrThrTyrProArgTyrGlyuThrAlaGlnThrThrSerGlyLysLeuThrGly 120
Db 435 GATTACACCTATCCGCGTTAGGAAACCAACCGCTGAACCAATCAGGGGTTTGAAGGGT 494
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db 495 TTAAACCACTTCTTTATCTTACACTTAATGCCCTGCACCTCGCGCACCAATCAGACCGT 554
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGly 160
Db 555 AGCGGAATGAGAGAGAGCTGGGCTTAAATATTGGCGGAGATGGGGATTTATCGAATATA 614
Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db 615 ACCTTGACGACCAACCCGCGGACACCTGCTTTTCCACTTGCTACAGCCGTAATTT 674
Qy 181 PheLeuAlaGlyIleLeuAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db 675 TTCCGCGCGGCAATAGAGGTGTTCTCTCGCAATGCGGATACGATGTGTTATTTAAC 734
Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHisLeuTyrAsnAlaGlnThr 220
Db 735 ATTCAGCATTTGCGAAGCATATGACGAAACCGAAATGCACCTTATCAATGCCAACA 794
Qy 221 LeuLysAlaGlnThrLysLeuGlyTyrPheAlaValAspArgThrAsnLysLysLeuLeu 240
Db 795 CTGAAGAGCCCAAACTGGAACTATTTCCGAGTAGACAGAACTAAATTAATTTGCTC 854
Qy 241 IleLysProLysThrAsnAlaPheGlnAlaAlaIleTyrLysGlnAspTyrAlaLeuTyrMet 260
```

```
Db 855 ATCAAAACCAAAACCAATGCCTTTGAAAGCTCCCTATAAAGAAATTAACGATTTGATG 914
Qy 261 GlyProTyrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer 280
Db 915 GGGCGGTATTAAGTAGAAGAAAGAAATCAACCGACGGAAGGATTTATGATTTCTCC 974
Qy 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGlnAlaAspAsn 300
Db 975 GATATCCGGCCATACGGGAATCATACGGGTAACTCCGCCCATCGTACAGGCTGATTAAC 1034
Qy 301 SerHisGlyGlyTyrGlyTyrSerAspGlnAlaValArgGlnHisArgGlnGlyGlnPro 320
Db 1035 AGTCATGAGGGGTATGATACAGGATGAAAGACAGTGCACAACTAGCAAGGGCAACT 1094

RESULT 2
US-10-617-835-6
; Sequence 6, Application US/10617835
; Publication No. US20050124037A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHUEBERPLUG, Ina
; APPLICANT: MAIER, Jurgen
; APPLICANT: EICKENRAGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; FILE REFERENCE: 0147-172P
; CURRENT APPLICATION NUMBER: US/10/617, 835
; PRIOR FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US/09/043, 302
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: PCT/EP96/04092
; PRIOR FILING DATE: 1995-09-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1094)
US-10-617-835-6

Alignment Scores:
Pred. No.: 3,21e-193 Length: 1136
Score: 1645.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-617-835-4 (1-320) x US-10-617-835-6 (1-1136)

Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db 135 ATGCGGCGACGGCTGCTGATACCTATCTTTTTCGTTTATTTATTCGCCCTGCGGG 194
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGlnGlnLeu 40
Db 195 ACACGACAGGATATTCATCGCATGCGAGGCAAACTTCGCGGTGACAACTT 254
Qy 41 ValAlaIaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60
Db 255 GTGGCGCTTCTGCGACAGCTCCGCTTAAAGACATGATTTACAGCATTTACAGGACGA 314
Qy 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Db 315 AAAGTTGATTTGATTCATTCGCAACTATGGGCGACCAAGTTTCAGGACGTTTGACAGGGGT 374
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
```

Db 375 CGCTACTCATTGATGACACTGATTCGGGGGAATACATAAGACCCCTGCGTCCGACAC 434  
 QY 101 AapTYrThTYrProARgTYrGLuThrThraIaGIuThrThSergIyGLyLeuThrgIy 120  
 Db 435 GATTACACCTATCCGGGTACGAAACCAACCGCTGAAACAAACATCAAGCGGTGTGACGGGT 494  
 QY 121 LeuThrThSergIyLeuThrThraIaProIaLeuSergIyThrgIInserAapGIy 140  
 Db 495 TTAAACCACTTCTTTATCTACCTTAATGCCCCCTGACCTCGCGACCCCAATCAGACGGT 554  
 QY 141 SergIySergSergSergLeuGLyLeuAenIleGIyGLyMeGlyAapTYrArgAngIu 160  
 Db 555 ACCGGAAGAGAGAGACGCTGCGCTTAATATTTGGGGATGGGGATTTATGAATGA 614  
 QY 161 ThrLeuThrThraAenProARgAphThraIaPheLeuSergIyLeuValGIuThrValPhe 180  
 Db 615 ACCTTGACACCAACCCCGCGACACCTGCTTTCTTCCACTGTGTACAGACCGATATT 674  
 QY 181 PheLeuARgGIyIleAapValIaSerProIaAenIaAapThraAapValPheIleAen 200  
 Db 675 TTCTCGCGGGCATAGACCTGTGTCTCTGCGCAATGCGATACAGATGTGTTATTAAAC 734  
 QY 201 IleAapValPheGIyThrIleARgAenARgThrgIuMeThIleuThyAenIaGIuThr 220  
 Db 735 ATCGACGATTTGGAACGATACGCAACGAAACGAAATGCACCTATACAAATGCCGAAACA 794  
 QY 221 LeuValaIaGIuThrThyLeuGLuThrPheIaIaValAapARgThraAenIyLeuLeu 240  
 Db 795 CTGAAGACCCCAACAAACCTGGAATATTTCCGAGTAGACAGAACCAATGAAATTTGCTC 854  
 QY 241 IleValaIaLeuThyThraAenIaPheGIuAaIaIaIyThyLeuGIuAenThyAaIaLeuThrPme 260  
 Db 855 ATCAAAACCAAAACCAATGCGTTGAGCTGCTTAAAGAAATTAATCCATTTGGAGT 914  
 QY 261 GIyProTYrLyVaIaSerIySGLyIleYsProThrgIuGLyLeuMeValaAapPheSer 280  
 Db 915 GGGCGGTATTAAGTAAGCAAAAGAAATCAAAACCGAAGGATTAATGTCATTTCTCC 974  
 QY 281 AapIleAapProTYrGIyAaenIeThrgIyAaenSerIaProSergValGIuIaAapAen 300  
 Db 975 GATTATCCGGCCATACGCGCAATCATACGGGTAACTCCGCCCATCCGTAGAGCTGATAC 1034  
 QY 301 SerIaGIuGLyThyGIyThySergAapGIuAaIaValaIaArgIInHIArGIuInGIuInPro 320  
 Db 1035 ACTCATGAGGGGTAAGATACAGCGATGAGGAGTGCGCAACATAGCAAGGGCACTT 1094  
 RESULT 3  
 US-09-043-302-1  
 ; Sequence 1, Application US/09043302  
 ; Publication No. US20020086349A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MEYER, Thomas F.  
 ; APPLICANT: RUDEL, Thomas  
 ; APPLICANT: SCHEUBERPLUG, Ina  
 ; APPLICANT: MAIER, Jurgen  
 ; APPLICANT: EICKERNJAGER, Sandra  
 ; APPLICANT: SCHWAN, Thomas  
 ; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
 ; FILE REFERENCE: 0147-172P  
 ; CURRENT APPLICATION NUMBER: US/09/043,302  
 ; EARLIER FILING DATE: 1998-06-08  
 ; EARLIER APPLICATION NUMBER: PCT/EP96/04092  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3287  
 ; TYPE: DNA  
 ; ORGANISM: Neisseria gonorrhoeae  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (136)..(447)

; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (583)..(1542)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1585)..(3111)  
 US-09-043-302-1  
 Alignment Scores:  
 Pred. No.: 1,636-192 Length: 3287  
 Score: 1645.00 Matches: 320  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-10-617-835-4 (1-320) x US-09-043-302-1 (1-3287)  
 QY 1 MetARgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGIy 20  
 Db 583 ATCGGGGACGGCTCTGATACCTATCTTTTTCGTTTATTATTATCCGCTCGGG 642  
 QY 21 ThrLeuThrgIyIleProSergIyGLyLyAaRgPheAlaValGIuInGIuLeu 40  
 Db 643 ACACAGACAGGATTCATTCGATCGACGCGGAGCAACGCTTCCGCTCGAACAAAGATT 702  
 QY 41 ValAlaIaSerAlaArgAlaAlaValIyAapMeAapLeuGIuAlaLeuHISGIyATG 60  
 Db 703 GTGGCGGCTTTCGACAGCTGCCCTGTAAAGCATGATTTTACAGCATTTACACGACGA 762  
 QY 61 LySValaIaLeuThyIleAlaThMeGIyAapGIuInGIySergIySergLeuThrgIyGIy 80  
 Db 763 AAAGTTGCATTTGATTCATTCGACCTATGCGGACCAAGGTTTCAGGACGTTTGCAGGGGT 822  
 QY 81 ArgTYrSerIleAapAlaLeuIleARgIyGIuThyIleAenSergProAlaValaIaArgThr 100  
 Db 823 CGCTACTCATTTGATGACCTGATTCGCGCGGAATACATTAACAGCCCTGCCGACAC 882  
 QY 101 AapTYrThTYrProARgTYrGLuThrThraIaGIuThrThSergIyGLyLeuThrgIy 120  
 Db 883 GATTACACCTATCCCGCTTACGAAACCAACCGTGAACAAACATCAAGCGGTTCAGCGGT 942  
 QY 121 LeuThrThSergIyLeuThrThraIaProIaLeuSergIyThrgIInserAapGIy 140  
 Db 943 TTAAACCACTTCTTTATCTACCTTAATGCCCCCTGACCTCGCGACCCCAATCAGACGGT 1002  
 QY 141 SergIySergSergSergLeuGLyLeuAenIleGIyGLyMeGlyAapTYrArgAngIu 160  
 Db 1003 AGCGGAAGTAGAGACAGCTGGGCTTAATATTTGGCGGATGGGGGATTAATGAAATGA 1062  
 QY 161 ThrLeuThrThraAenProARgAphThraIaPheLeuSergIyLeuValGIuThrValPhe 180  
 Db 1063 ACCTTGACACCAACCCCGCGACACCTGCTTTCTTCCACTTGTACAGACCGTAAATT 1122  
 QY 181 PheLeuARgGIyIleAapValIaSerProIaAenIaAapThraAapValPheIleAen 200  
 Db 1123 TTCTCGCGGGCATAGACCTGTGTCTCTGCGCAATGCGGATACGATGTGTTATTAAAC 1182  
 QY 201 IleAapValPheGIyThrIleARgAenARgThrgIuMeThIleuThyAaIaLeuThr 220  
 Db 1183 ATCGACGATTCGGAACGATACGCAACGAAACCGAAATGCACTTATACATGCCGAAACA 1242  
 QY 221 LeuValaIaGIuThrThyLeuGLuThrPheIaIaValaIaAapARgThraAenIyLeuLeu 240  
 Db 1243 CTGAAGACCCCAACAAACCTGGAATATTTCCGAGTAGACAGAACCAATTAATAAATTTGCTC 1302  
 QY 241 IleYsProLYsThraAenIaPheGIuAaIaIaIyThyLeuGIuAenThyAaIaLeuThrPme 260  
 Db 1303 ATCAAAACCAAAACCAATGCGTTTGAACCTGCTTAAGAAATTAAGCAATTTAGCATTTGGAGT 1362  
 QY 261 GIyProTYrLyVaIaSerIySGLyIleYsProThrgIuGLyLeuMeValaAapPheSer 280  
 Db 1363 GGGCGGTATTAAGTAAGCAAAAGAAATCAAAACCGAAGGATTAATGTCATTTCTTCC 1422

```
OY 281 AsplleargProTYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300
|||
DB 1423 GATATCCGGCCATACGGCAATCATTCGGGTAACTCCGCCCATCCGTTAGAGGCTGATTAAC 1482
OY 301 SerHisGluGlyTYrGlyTYrSerAspGluAlaValArgInHisArgGlnGlyGlnPro 320
|||
DB 1483 AGTCATGAGGGGTATGATGATACGCGATGAGCAGTGCACAAACATGACAAAGGGCAACTT 1542

RESULT 4
US-10-617-835-1
; Sequence 1, Application US/10617835
; Publication No. US20050124037A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHUEBERPLOG, Ina
; APPLICANT: MAIER, Jurgen
; APPLICANT: EICKERMJAGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; TITLE OF INVENTION: the Adhesion of Neisseria Cells to Human Cells
; FILE REFERENCE: 0147-172P
; CURRENT APPLICATION NUMBER: US/10/617, 835
; PRIOR FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US/09/043,302
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: PCT/EP96/04092
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3287
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(447)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (583)..(1542)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1585)..(3111)
US-10-617-835-1

Alignment Scores:
Pred. No.: 1,63e-192 Length: 3287
Score: 1645.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-617-835-4 (1-320) x US-10-617-835-1 (1-3287)
OY 1 MetaArgAlaArgLeuLeuIleProIleLeuPheIleLeuSerAlaCysGly 20
|||
DB 583 ATGGGGGACGCGCTGCTGATACCTATCTTTTTCGGTTTATTTATTCGCCGCTGGGG 642
OY 21 ThrLeuThrGlyIleProSerHisGlyGlyGlyIleArgPheAlaGluGlnGluLeu 40
|||
DB 643 ACACGTGACAGGTATTCATCCATGCGAGCGAGCAAAACCTTCGCGGTGGAACAAGAACTT 702
OY 41 ValAlaAlaSerAlaArgAlaAlaValIleAspMetAspLeuGlnAlaLeuHisGlyValArg 60
|||
DB 703 GTGGCCCGTTGTGCGAGGCTGCCGTTAAAGACATGATGATTTACAGGCACTTACAGGACGA 762
OY 61 LysValAlaLeuTYrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
|||
DB 763 AAAGTTGATTTGATTCATTCGCAACTRATGCGACCAAGGTTTCAGGACGTTTACACAGGGGT 822
OY 81 ArgTYrSerIleAspAlaLeuIleArgGlyGlyIleAsnSerProAlaValArgThr 100
```

```
DB 823 CGCTACTCCATTGATGACCTGATTCGGCGGCAATATACATAAACAGCCCTCGCGCACCC 882
OY 101 AspyrThrTYrProArgTYrGluThrThrAlaGluThrThrSerGlyIleuThrGly 120
|||
DB 883 GATTACACCTATCCGCGTTAGAAACCAACCGCTTAACCAACATACAGGCGGTTTACCGGGT 942
OY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerAlaGlyThrGlnSerAspGly 140
|||
DB 943 TTAAACACTTCTTTATCATCACTTAATGCCCTTCCTGCACTCTGCCACCCATCAGACGGT 1002
OY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyIleMetGlyAspTYrArgAsnGlu 160
|||
DB 1003 AGCGAAGTAGAGACAGCTGCTGGCTTAAATATTGGCGGAGTGGGAGATTATCGAAATGAA 1062
OY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
|||
DB 1063 ACCTTGACGACCAACCCGCGGACACCTGCTTCTTCCCACTTGATACAGACCGTATTT 1122
OY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
|||
DB 1123 TTCCTGGCGGCGCATAGAGTTGTTCTCTGCGCAATGCCATACAGATGTTTATTATTAAC 1182
OY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTYrAsnAlaGluThr 220
|||
DB 1183 ATCGACGTAATCGGAACGATACGCACAGAACCGAAATGCACTTATCAATGCCGAAACA 1242
OY 221 LeuIleAspGlnThrIleAsnGluTYrPheAlaValAspArgThrAsnIleValLeuLeu 240
|||
DB 1243 CTGAAGGCCCAACCAACCTGGAATATTTCCAGATGACACAACTAAATTAATTTGCTC 1302
OY 241 IleIleAspProIleThrAsnAlaPheGluAlaAlaIleTYrIleGluAsnTYrAlaLeuTrpMet 260
|||
DB 1303 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTCTTAAGAAATTAACGCACTTGATGATG 1362
OY 261 GlyProTYrIleValSerIleGlyIleIleArgProThrGlnGlyLeuMetValAspPheSer 280
|||
DB 1363 GGGCGGTATTAAGTAAAGTAAAGGAATCAAAACCGACGGAAGATTAAATGTCGATTTCTCC 1422
OY 281 AsplleargProTYrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300
|||
DB 1423 GATATCCGGCCATACGGCAATCATTCGGGTAACTCCGCCCATCCGTTAGAGGCTGATTAAC 1482
OY 301 SerHisGluGlyTYrGlyTYrSerAspGluAlaValArgInHisArgGlnGlyGlnPro 320
|||
DB 1483 AGTCATGAGGGGTATGATGATACGCGATGAGCAGTGCACAAACATGACAAAGGGCAACTT 1542

RESULT 5
US-10-320-800-69
; Sequence 69, Application US/10320800
; Publication No. US20030215469A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, ANDREW
; APPLICANT: GORRINGE, ANDREW
; APPLICANT: HUDSON, MICHAEL
; APPLICANT: REDDIN, KAREN
; TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE
; FILE REFERENCE: 1581.0790001
; CURRENT APPLICATION NUMBER: US/10/320, 800
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: PCT/GB99/03626
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(939)
; OTHER INFORMATION: NMB0652 maf A
US-10-320-800-69
```

## Alignment Scores:

Pred. No.: 5,836-117 Length: 939  
 Score: 1029.00 Matches: 197  
 Percent Similarity: 78.34% Conservative: 49  
 Best Local Similarity: 62.74% Mismatches: 62  
 Query Match: 62.55% Indels: 6  
 DB: 17 Gaps: 2

US-10-617-835-4 (1-320) x US-10-320-800-69 (1-939)

```

QY 5 LeuIleuIleProIleLeuPheSerValPheIleLeuSerAlaCyGlyThrLeuThrGly 24
DB 16 CTCCTCTACGCTCC-----GTCCTTACGCTCCGCGCACCTGACCTGACGCGC 60
QY 25 ILeProSerHieGlyGlyValYsArgPheAlaValGlnGlnIleuValAlaIleSer 44
DB 61 ATACCGCGCCACGCGCGCGCAACGCTTGGCGTGAACAAAGAACTCGTCCGCGCATG 120
QY 45 AlaArgAlaAlaValYsArgPheSerValPheIleuHieGlyValYsValAlaIleu 64
DB 121 TCCCGCGCGCGCGCTCAAAAGAAATGGATTGTCCGCGCTAAAGAGACGCAAAACCGCGCTT 180
QY 65 TyrIleAlaThrMerGlyValArgGlnIleYsArgIleuThrGlyValArgTyrSerIle 84
DB 181 TACGCTCTCGGTATGGCGACCAAGGTTGGGCAACATAGCGCGGACGCTATCTATC 240
QY 85 AspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThrAspTyrThrTyr 104
DB 241 GACGCACTGATACGGCGCGGCTTACCAACAAACCCGAAAGTGGCCACCAATACAGCTAC 300
QY 105 ProArgTyrGlnThrThrAlaGlnThrThrSerGlyGlyLeuThrGlyLeuThrThrSer 124
DB 301 CCGCGCTACGACACTACCCGCAACCAATCCGACGCGCTCTCCAGCGTAAACCACTTCC 360
QY 125 LeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspIleYsArgIleuArg 144
DB 361 ACATGCGCTTTGAAAGCGCGCGCGCGCTGACGCAAAACGCGGACGCAAAAGCGCA 420
QY 145 SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlnThrLeuThrThr 164
DB 421 CGCTCGCGCGGACTGTCGCTCAACGCGGCGGAGCTACCGCAAGAAACCTGCTCGCGC 480
QY 165 AsnProArgAspThrAlaPheLeuSerHieValGlnThrValPhePheLeuArgGly 184
DB 481 AACCCCGCGGACGTTCTCTCTGACCAACCTCATCAAAACCGTCTTCACTGCGCGCGC 540
QY 185 IleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPhe 204
DB 541 ATCGAAAGTGTACCGCGCGCAATACGCGGACCGAGCTATGTGTAACGCGTATTC 600
QY 205 GlyThrIleArgAsnArgThrGlnMetHieValYsArgAlaGlnThrLeuIleuValGln 224
DB 601 GGCACGCGTCCGCGGCGGCTACCGCACTGCACTCTTAAACGCGGAAACCTTAAAGCGCA 660
QY 225 ThrIleValGlnTyrPheAlaValAspArgThrAsnIleYsValLeuIleuIleuPhePhe 244
DB 661 ACCAAGCTCGAATATTTCCGCGTTCGACGCGGCAACGCGGAAACGCTGTGATTCACCTTAA 720
QY 245 ThrAsnAlaPheGlnAlaIleAlaTyrIleValGlnIleuTyrAlaLeuThrMerGlyProTyrIle 264
DB 721 ACCGCGCGCTTACGAAATCCCAATACCAAGAAACATACGCGCTTGAACGCGCGCTTAA 780
QY 265 ValSerIleGlyIleValProThrGlnGlyLeuMetValAspPheSerAspIleArgPro 284
DB 781 GTACAGAAACCGTCAAGCGCTGACGCGCTGATGTGATTTCTCCGACATTAACCGCC 840
QY 285 TyrGlyAsnHieThrGlyAsnSerAlaProSerValGlnAlaAspAsnSerHieGlnGly 304
DB 841 TACGGGACAAACCGCGCAACCGTCCGACTTCAAAACAAACGATTAACAAACCC 900
QY 305 TyrGlyTyrSerAspGlnIleValArgGlnHieArgGlnGly 318
  
```

DB 901 ---GATGTCGCAAGAACTCATCCGCGCGCGCAAGAGCA 939

```

RESULT 6
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIORITY FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1
  
```

## Alignment Scores:

Pred. No.: 1,316+04 Length: 9025608  
 Score: 107.50 Matches: 50  
 Percent Similarity: 42.47% Conservative: 29  
 Best Local Similarity: 26.88% Mismatches: 52  
 Query Match: 6.53% Indels: 55  
 DB: 15 Gaps: 10

US-10-617-835-4 (1-320) x US-10-156-761-1 (1-9025608)

```

QY 6 LeuIleProIleLeuPheSerVal-----PheIleLeuSerAlaCyGlyThrLeuThr 23
DB 8122195 CTGTGCGCGTGTGTGCGCGCGTGGCGCGCGTGAACCGCGCGCTGCCACCGTGGCG 8122136
QY 24 GlyIleProSerHieGlyGly-----GlyIleValArg----- 33
DB 8122135 GGCACCCCAAGCGGAGACGCGCGGCTAGCGCTGCGGCGCGCGCGCGCGCGCGCG 8122076
QY 34 -----PheAlaValGlnIleu 40
DB 8122075 CTGCGACCACTGGAGCTCCAGCTGGGTAGCGACTGCACTGACGCTGCGGTATTCAGGAAATAT 8122016
QY 41 ValAlaIleSerAlaArgAlaAlaValYsArgPheAsp----- 53
DB 8122015 CCGCGGCGAGCC-----GCTGCACTGCTGCAATCGAGCACTTGGCGGTATGTGTG 8121962
QY 54 -----LeuGlnAlaLeuHieGlyArgGlyValAlaLeuTyrIleAlaThr 68
DB 8121961 CATCGCGCGATGGCACTTCGACATGATGTCG-----CTCAACTCTTCGCGCG 8121911
QY 69 MetGlyAspGlnGlySerGlySerLeuThrGlyGlyArgTyrSerIleAspAlaLeuIle 88
DB 8121910 GTCGGTGTGTCTTCCGCGAGTGAACAACCTGCGCGCGGTGTGTGTGCACTTAACG 8121851
QY 89 ArgGlyGlnTyrIleAsnSer-----ProAlaValArgThrAspTyrThrTyrPro 105
DB 8121850 GCACCTCGGTGATGAACCTCGCGCGGACTCGGCCATTCGATCGCAGCGCGTCAATC--- 8121794
QY 106 ArgTyrGlnThrThrAlaGlnThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeu 125
  
```

Db 8121793 -----ACGGCTGCTCCACCTGCTGTCACAGGAGCGTGGCGGTC---ACGGCTTG 8121746  
Qy 126 SerThrLeuAsnAlaPro-----AlaLeuSerArgThrGlnSerAspGly 140  
Db 8121745 GCGACGGCCACACGCGCGCTGTCACGCGTGAGAGCGCGCGCGCGCATCCAGCACGGGC 8121686  
Qy 141 SerGlySerArgSerSer 146  
Db 8121685 CAGGGGTGCGCGCTGCGCG 8121668  
  
RESULT 7  
US-10-398-221-10  
; Sequence 10, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederick  
; APPLICANT: GLASER, Philippe  
; TITLE OF INVENTION: Listeria innocua, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 1163020  
; TYPE: DNA  
; ORGANISM: Listeria innocua  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(end)  
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u  
US-10-398-221-10  
  
Alignment Scores:  
Pred. No.: 1.35e+03 Length: 1163020  
Score: 104.50 Matches: 77  
Percent Similarity: 35.24% Conservative: 40  
Best Local Similarity: 23.19% Mismatches: 113  
Query Match: 6.35% Indels: 102  
DB: 17 Gaps: 16  
US-10-617-835-4 (1-320) x US-10-398-221-10 (1-1163020)  
  
Qy 9 lIleuPheSerValPheIleuSerAlaCys-GlyThrLeuThrGlyIleProSerH 28  
Db 62741 ATTTATATCTGAGGCTGGAATTAAGTAGCTGCGGGAACAGACACTGCATTGCCAGA-- 62798  
Qy 28 sGlyGlyGlyAsnArgPheAlaValGlnGlnLeuValAlaIleSerAlaArgAla 48  
Db 62799 -----CTTTCCAAAGCGCTGAAAAAC 62821  
Qy 48 aValIleAspMetCAspLeuGlnAlaLeuHisGlyArgIleValAlaLeuThrIleAla 68  
Db 62822 GAATTAAGATGCAATGCTGCTGCTCATTATGCGGCAACAATGAACCAATTAAC 62881  
Qy 68 rMetClyAspGlnGlySerGlySerLeuThrGlyIleArgTyr-----SerIleAs 85  
Db 62882 T-----GGTGTATGGGGCGCGTGTGTGAGATTGGCGGTATAGAGATTGA 62929  
Qy 85 pAlaLeu-----IleArgGlyG 91  
Db 62930 TGGCTTAATAAGCAAGACCCCACTTTAAAGATATTGGATGATCTGCCAAAGCGCC 62989  
Qy 91 uTyrIleAsnSerProAlaValArgThrAspTyrThrTyrProArgTyrGluThrThrAl 111  
Db 62990 GTATCTTATATAGTAAGAACTGGGAGATTATACATAT----- 63029  
Qy 111 aGluThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuAsnAlaPr 131

Db 63029 ----- 63029  
Qy 131 oAlaLeuSerArgThrGlnSerAspGlySerArgSerLeuGlyLeuAsnI 151  
Db 63030 -----GACAAATCGGCTTCAATTGCTATAGGCACAGATGTAAAGTAATTAATCT 63082  
Qy 151 eGlyGlyMetClyAspTyr-----ArgAsnGluThrLeuThrThrAsnProArgAs 168  
Db 63083 AGGTGGATGATTTCTTGATGACGAGTCGACAGATAAACACGACTTCACATAGCGCGA 63142  
Qy 168 pThrAlaPheLeuSerHisLeuValGlnThrValPhePheLeuArgGlyIleAspValVa 188  
Db 63143 TGA-----CTTCAAAAGCTATTATAAACCGGATGTTCT-----GGTACAGTCGAT 63190  
Qy 188 lSerProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleAr 208  
Db 63191 C---CCGAAACCAATCA-----TATGCCAATTATATGTGTAGCACTGTAA 63241  
Qy 208 gAsnArgThrGlnMetHisLeuTyrAsnAlaGluThrLeuValAlaGlnThrIleVal 228  
Db 63242 ACCTTATATAGTAATAAGAGTCGATACGAAATTACGATTACTATATGAAAA----- 63296  
Qy 228 uTyrPheAlaValAspArgThrAsnIleValLeuLeuIleValProIleThrAsnAlaPh 248  
Db 63297 -----GCGGATGAACCAATGA-----GTCTTGAATCAACAGAGCTATCTT 63340  
Qy 248 eGluAlaAla-----TyrIleGluAsnTyrAlaLeuTrpMe 260  
Db 63341 TGAACAGTAAATTTACCTTAAGTTTATTTATTCAGTAAAGCCGCGCAACGTTACTGC 63400  
Qy 260 rGlyProTyrIleValSerIleValGlyIleValProThrGluGly-----LeuMetValAs 278  
Db 63401 TGAAGATTATTAACA-----GGTACGGTTACTACTAGGTGCAATACAGATGTGA 63454  
Qy 278 pPheSerAspIle-----ArgProTyrGlyAsnHisTh 289  
Db 63455 TTTAGCTTGGGTGTATGATGACACGAATAATCCGCAAGCGCATCTTACTTCCGCTT 63514  
Qy 289 rGlyAsnSerAlaProSerValGluAlaAspAsn 300  
Db 63515 AAAATCGATGCACTGTGTGTGATGTAAATAT 63548  
  
RESULT 8  
US-10-398-221-2058  
; Sequence 2058, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederick  
; APPLICANT: GLASER, Philippe  
; TITLE OF INVENTION: Listeria innocua, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2058  
; LENGTH: 3011208  
; TYPE: DNA  
; ORGANISM: Listeria innocua  
US-10-398-221-2058  
  
Alignment Scores:  
Pred. No.: 5.78e+03 Length: 3011208  
Score: 104.50 Matches: 77  
Percent Similarity: 35.24% Conservative: 40  
Best Local Similarity: 23.19% Mismatches: 113  
Query Match: 6.35% Indels: 102  
DB: 17 Gaps: 16

```

US-10-617-835-4 (1-320) x US-10-398-221-2058 (1-3011208)
QY      9  IleuPheSerValPheIleuSerAlaCys-GlyThrLeuThrGlyIleProSerHi 28
      146982 ATTTATACCTGAGGCTGGATAAAGTAGCTCCGGAAACAGACACTGATCCAGAGA-- 147039
QY      28  GcIyGlyGlyValArgPheAlaValGluGluIleuValAlaIleSerAlaArgAlaAl 48
      147040 -----CTTTCCAAAGCGGCTGAAAAAAC 147062
QY      48  aValIySaPMeCaPLeuGluAlaLeuHiGcIyArgValAlaLeuTyrIleAlaTh 68
      147063 GAATTAAGATGAGTGGCTGCTCCTATGCGGAAACATGAAACCCCAATTAAAC 147122
QY      68  tMeCgIySaPngIySerGlySerLeuThrGlyGlyArgTyr-----SerIleAs 85
      147123 T-----GGTGAATGGGGCGGCTGCTGATGAGTGGCGGTATAGAAATATGA 147170
QY      85  PAlaLeu-----IleArgGlyGly 91
      147171 TGCTCTAAAGCCCAAGACCCCACTTTAAAGAAATATGGGATGATCTGCCAAAGCGCC 147230
QY      91  uTyrIleAsnSerProAlaValAlaArgThrAspTyrThrTyrProArgTyrGluThrThrAl 111
      147231 GATCTTTATAGTAAGAAACAGTGGGAGTTTACACATAT----- 147270
QY      111  aGluThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuAsnAlaArg 131
      147270 ----- 147270
QY      131  cAlaLeuSerArgThrGlnSerThrAspGlySerGlySerArgSerLeuGlyLeuAsnAl 151
      147271 -----GACAAATACCGCTTCAATGCTATAGACACGATGTAAAGATTAATATCT 147323
QY      151  eGcIyGlyMeCgIyAspTyr-----ArgAsnGluThrLeuThrThrAsnProArgAs 168
      147324 AGGTGGATGATTTCTTGATGATCAGTCGAGATTAACACGACTTCACTCAACGCGCA 147383
QY      168  pThrAlaPheLeuSerHiLeuValGlnThrValPhePheLeuArgGlyIleAspValVa 188
      147384 TGAA-----CTTACAAAGGCTATTAAACCGGATTTGTC-----GGTACAAAGTCAT 147431
QY      188  lSerProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleAr 208
      147432 C---CCGCAAAACGCAATCACA---TATCGAATTTAAATGTGTAGCAACTGTAA 147482
QY      208  gAsnArgThrGluMeCHisLeuTyrAsnAlaGluThrLeuValAlaGlnThrTyrLeuGl 228
      147483 ACCTTATAGTGAATAATGAGTCGATACGAATTCGATTAATATGAABAA----- 147537
QY      228  uTyrPheAlaValaAspArgThrAsnIyLeuLeuIleLeuProLysThrAsnAlaArg 248
      147538 -----GCGGATGAAACAAATGAA-----GTCTGAAATCAACAGACCTATCTTT 147581
QY      248  eGluAlaAla-----TyrLysGluAsnTyrAlaLeuTyrMe 260
      147582 TGAACAGTAAATTAACCTAAGTTTATATTCAGTAAAGCGCGAAACGCTTACGTC 147641
QY      260  cGcIyProTyrValSerIyGcIyIleLysProThrGluGly-----LeuMeCValAs 278
      147642 TGGAAATTAATAAGCA-----GTTACGCTTACTTCACTGCGCAATACAGTGTGA 147695
QY      278  pPheSerAspIle-----ArgProTyrGlyAsnHiThr 289
      147696 TTATACCTCGGTATGATGACACAGCAATTCGCAAGCGCATCTTATATCTCCGCTT 147755
QY      289  cGcIyAsnSerAlaProSerValGluAlaAspAsn 300
      147756 AAAATCGAGTCACTTATGTGTGATTAATAT 147789

```

```

Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Bardazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 87357
LENGTH: 2604
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: clone ID: PAT_MRT4530_8630C.1
US-10-437-963-87357
Alignment Scores:
Pred. No.: 0.161 Length: 2604
Score: 103.50 Matches: 86
Percent Similarity: 34.66% Conservative: 53
Best Local Similarity: 21.45% Mismatches: 129
Query Match: 6.29% Indels: 133
DB: 19 Gaps: 20
US-10-617-835-4 (1-320) x US-10-437-963-87357 (1-2604)
QY      19  CysGly--ThrLeuThrGlyIleProSerHiGcIyGlyLys-----ArgPhe 34
      274 TGTGCTTGTCCTTATGATGAGGTACACGATACCGGACCGACAGAGGATTAATCTCAGGTTCC 333
QY      35  AlavalGluGlnGluLeuVal----- 41
      334 GCCGACCCGACGCCGCGCATGCTGCCAGATGAGCGCGGACCCAGACTCCGCCCTTCC 393
QY      42  AlaAlaSerAlaArgAlaAlaVal----- 49
      394 TCCACGCGTCGAGGCGCGCTCAGCTCGGGGACCGAACACCAAGAGAGCTGGAGTTCC 453
QY      50  -----LysAspMeCAspLeu 54
      454 ATCTCGGACGAGTGCCTCACTATATACGCGCATGCTCATGAGAGAGAGACTTGAT--- 510
QY      55  GluAlaLeuHiGcIyArgLysValAlaLeuTyrIleAlaThrMeCgIyAspGlnLys 74
      511 -----GAGAAAGCTCAGATGATAC-----CAGAGAGAGTCC 540
QY      75  GlySerLeu-----ThrGlyLysArgTyr----- 82
      541 GCGGCGCTCGCGGCTGCTGCGAAGCATTTATGACATTTCTGGACACAGTTCCCGCC 600
QY      83  SerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValAlaArgThrAspTyr 102
      601 TCCCTGACCGGACGAGCTGTAGCTTGCGCCCTCGACACCGGAGTGAAGTACACTGAC 660
QY      103  ThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThr 122
      661 AGCTATCA---CACTCTTGAGTATGTTCTGTATACACACAACTATGAGTGTCCGCTC 717
QY      123  ThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142
      718 GATAC-----ACCAACCGCTTATGTTGACACAGT 750
QY      143  SerArgSerSerLeuGlyLeuAsn-----IleGlyIleMeCgIyAspTyr 157
      751 GAGTATCGAGCTTGCTGTGATCTTCTCAACCGCGGTGGCCCATCAAGTATGTT 810

```

```

QY 158 ArgAsn-----GluThrLeuThrAspProArgAspThrAlaPhe 171
      |||
      |||
      |||
Db 811 CGCAACGCTATGAGACACTAGAAATCCTTTGATTAGTAATGGACAGATCCCTGATAT 870
QY 172 LeuSerHisLeuValGlnThrValPhePheLeuArgGlyIleAspValValSerProAla 191
      |||
      |||
      |||
Db 871 TTGTTTGAGAGCTTTCACACATGGGATTTCAAGAGAGGTGTGACAGAGCACAAGTTT 930
QY 192 AsnAlaAspThrAsp---ValPheIleAsnIleAspValPheGlyThrIleArgAsnArg 210
      |||
      |||
      |||
Db 931 CTCGCCGGGTAGCGATGAGTACTGATGATTTAGAGGCTGTGGTGTCGAACAGCTCA 990
QY 211 -----ThrGluMetHisLeuTyIAsnAlaGluThrLeuIleValGln--- 224
      |||
      |||
      |||
Db 991 GAAGCAGGGAAGAGCTATTTCTGTGAATGTCAGCAAGCAGAGGTTTGAAGGTAAAGAAA 1050
QY 225 -----ThrIleLeuGluIleTyIAsnValPhe 230
      |||
      |||
      |||
Db 1051 AACAGACAGAGTGAAGACCTTGACGTGATGGAAGACGGAATATGTAACAGTCTGCATTT 1110
QY 231 AlaValAsp-----ArgThrAsnIle 237
      |||
      |||
      |||
Db 1111 TGTTCCTGATGAACCTGATTTGATGATGTTTGTCTGCTTCTGCTCAACGTAGAG 1170
QY 238 LysLeuLeuIleLysProLysThrAsnAlaPheGluAlaAlaTyIArgLysGluAsnTyIAla 257
      |||
      |||
      |||
Db 1171 AAGGCAACGTGATCTTGAAAAATGATGCGTTTGGAGGCTTCCAAAGATTTCCGCCCTTGT 1230
QY 258 LeuTrpMetGlyProTyIAsnVal---SerLysGlyIleLysProThrGluGlyLeuMet 276
      |||
      |||
      |||
Db 1231 CAGCCAAAAGACCAAGTGGGACCCGACCGGACCGGAGCGGAGCCGAACTTAAG----- 1281
QY 277 ValAspPheSerAspIleArgProTyIAsnHisThrGlyAsnSerAlaProSerVal 296
      |||
      |||
      |||
Db 1282 AAGGATGGGTGGACCTTAGACGCTACTCATCCAC-----TGGCTCAGGCAAG 1332
QY 297 GluAlaAspAsnSerHisGluGlyTyIArgSerAspGluAlaValArgGlnHisArg 316
      |||
      |||
      |||
Db 1333 GCAGCTGATGACCGCGGAGCT-----GCTAATGAGTGTCTTAAGCAATAAG 1380
QY 317 Gln 317
      |||
      |||
      |||
Db 1381 CAG 1383

RESULT 10
US-10-282-122A-16487
; Sequence 16487, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Twick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335

```

```

; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 16487
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Clostridium botulinum
US-10-282-122A-16487

Alignment Scores:
Pred. No.: 0.118 Length: 1335
Score: 101.00 Matches: 71
Percent Similarity: 35.99% Conservative: 42
Best Local Similarity: 22.61% Mismatches: 111
Query Match: 6.14% Indels: 90
DB: 17 Gaps: 15

US-10-617-835-4 (1-320) x US-10-282-122A-16487 (1-1335)
QY 62 ValAlaLeuTyIleAlaThrMetGlyAspGlnGlySerGly----- 75
      |||
      |||
      |||
Db 133 ATACCTTTTATGATATCTTAATA-----GATTCGAGATTCGACTATATATA 183
QY 76 SerLeuThrGlyArgTyIAsnValAspAlaLeuIleArgGlyIleAsnSer 95
      |||
      |||
      |||
Db 184 TCTGTAACTGAAGCAGAAAGCAATATAGCTTGTGTAATAAGATATTAAACATA 243
QY 96 ProAlaValArgThrAspTyIThrTyIProArgTyIArgIleThrAlaGlnThrSer 115
      |||
      |||
      |||
Db 244 TCCTTAGATTCTAAAAAAGTAAATTAATATATATAAAAAAGATTATTAGTTAAGGGT 303
QY 116 -----GlyGlyLeuThrGlyLeuThrSerLeuSer----- 126
      |||
      |||
      |||
Db 304 CTTAAAAATGTTGCTATAGGTGAGAGAACTGGCTTATCTATGTTAAAGACATTAAA 363
QY 127 -----ThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySer--- 143
      |||
      |||
      |||
Db 364 TATTATACCTCTAATATTAACGCTATGTTACAGTTGCAGATGATGGGGATCTGAT 423
QY 144 -----ArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyIArgAsnGlnThr 161
      |||
      |||
      |||
Db 424 GAATTAAGAGAAAGATTATGACATGTA-----CTTCAGAGAGATATTAAGAACTGTATA 477
QY 162 LeuThrThrAspProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe--- 180
      |||
      |||
      |||
Db 478 TTATCTTTATCAGATACAGAA---CCTTTAATGGAAGAACTTTTACATATAGATTACA 534
QY 181 -----PheLeuArgGlyIleAspVal 188
      |||
      |||
      |||
Db 535 GATGGAAGACTTAATAAACCAAGTTTGTAAATTATTTTACGTGATGATGAGATA 594
QY 189 Ser----- 189
      |||
      |||
      |||
Db 595 TCTAATATTTGAGGAAGAGTTCAAAAAGTTAGTTCGGTTTGGCTGTACAGAAAA 654
QY 190 -----ProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIle 207
      |||
      |||
      |||
Db 655 GTTGTAACCACTTAAGAG-----AAATATAGTGTCTAAAGCTAAGGCTG 699
QY 208 ArgAsnArgThr-----GluMetHisLeuTyIAsnAlaGluThrIleLeuValAla 223

```



```
Dbb      700 AAAAAACATGATAGTAGAGGAGAAATCCAAATATA---CCAGAAAAGAATCTTCACTAT   756
Oy      224 GlnThrLysLeuGluTy rPheAlaValAspArgThrsnLysLysLeuIleLysPro  243
Dbb      757 AATAGGTAAATATGAGAGMGGTATTATTATAGAGCCGAAGAAATCTCAAGGCACATCATGAGGCT  816
Oy      244 LysThrAsnAlaPheGluIlaIaIeTy rLVYgLuAsnTy rAlaIeuTrpMetGlyPro--  265
Dbb      817 GTTACA-----GCTAATAAAGAAAGCAGATCCCATCATCATTTGGCGCTGGA  861
Oy      263 -----TyrLy sValSerLy sGlyIleLys  270
Dbb      862 AGCTTTTATACCAGCGCTAATACTTAATTTGTATAATAAAGATATCACAGAAAGCTTTAAA  921
Oy      271 ProThrGluGlyLeuMetVala sPheSerAspIleArgProTy rGIa s nH i s t h G l y  290
Dbb      932 AAGACAAAGCGTCCTAAATATATATATATACAAATATATAGCATCAACCCAGAGAGAAACAGAT  981
Oy      291 AsnSerAlaProSer-----ValGIuaIla sP asSerHis  302
Dbb      982 AATTTCATCTTGTCAGATCATATAAAAACATATAAATTAACAT  1023
```

```

RESULT 11
US-10-425-114-12803
; Sequence 12803, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; TITLE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 12803
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 701212455_FLI
US-10-425-114-12803

```

Alignment Scores:	
Pred. No.:	0.186
Score:	101.00
Percent Similarity:	40.68%
Best Local Similarity:	24.86%
Query Match:	6.14%
DB:	18
Gaps:	6
US-10-617-835-4 (1-320) × US-10-425-114-12803 (1-1795)	

```
OY      20 G|T|T|h|e|t|h|G|L|I|e|r|o|s|e|r|h|S|G|I|G|L|y|L|y|a|r|g|---P|e|a|V|A|L|G|I|u|g|I|n   38
Db      216 GGAACCTTCAGAAACGCTTCTGTGGTGAGAGCGTCGACGANGAACCAAGCCTGTCCAAAGCCT    275
OY      39 G|U|E|V|A|I|A|L|A|S|e|r|A|I|a|r|g|A|A|A|V|A|L|y|A|S|P|m|e|r|A|B|r|e|U|G|I|N|A|L|e|U|h|S   58
Db      276 CCATTATATACGACAACCTTCTTCTGCT-----                               305
OY      59 G|I|A|r|g|L|y|V|A|L|A|L|e|U|r|y|r|I|e|A|I|a|r|m|e|r|G|I|A|S|p|c|i|n|l|y|S|e|r|g|I|S|e|r|U|t|h|r   78
Db      306 -----GCCATTGACACTGGCTGCTTCATAATATTTCTTCAACTCTTCTTGCT    353
OY      79 G|I|G|I|A|r|g|T|r|S|e|r|I|e|A|P|A|L|e|U|I|e|A|r|g|I|G|U|r|y|r|I|e|A|n|S|e|r|-----          95
Db      334 CCACTACAAACCCCCCATGCACACTGCTTCAACCTCAAAATCACATCAAACCTTTGTTTTAT    413
```

```

QY 96 -----ProAlaValArgThrAspPylrThrYrProArgTgclunThrAlaIgu--- 112
Db 414 GGCTTACCAAGTACGCTTACGCATGTGATCTTCCATTCTCTCAAGTTCGGTCTTAGA 473
QY 113 ---ThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThr---LeuAlaAla 130
Db 474 ATCTCCACTCCGGGTTTGATCTTCAGCTGAATTAATGCCCTTGATTAAGTATTTTATCC 533
QY 131 ProAlaLeuSerArgThrGlnSerAspGlySerGlySerArgSerSerLeuGlyLeuAen 150
Db 534 GGGGTTTGTCCAAAGAAAGCTAGTATATATACGGTACAGAAACGTTGGTTCCAAAC 593
QY 151 IleGlyGlyMetGlyAspTyrArgAngIunThrLeuThrThrAsnProArgAspThrAla 170
Db 594 AATACGCTTCTTTCAAAGTTATATACATCCACACTAGTACTACTACCTACCTATGTCTTC 653
QY 171 PheLeu---SerIleLeuValGlnThrValPhePheLeuArgGlyLeuAsp 186
Db 654 CTTTGGACTCTTCTTGGCTTCAACAGAAATTCATGACAGATGGTGTGAT 704

```

```

RESULT 12
US-10-424-599-56736
: Sequence 56736, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53423)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ. ID NOS: 285684
: SEQ. ID NO 56736
: LENGTH: 1946
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_22243C.1
US-10-424-599-56736

```

Alignment Scores:	
Pred. No.:	0.211
Score:	101.00
Percent Similarity:	40.66%
Best Local Similarity:	21.86%
Query Match:	6.14%
DB:	18
Gaps:	6
Length:	1946
Matches:	44
Conservative:	28
Mismatches:	81
Indels:	24

Oy	20	GLYTHRLEUTHRGLYLIEPROSERHISGLIYGLIYLSARG---	PHEALVALGLUGLN	38
Db	367	GGAATCTTCAGAAACGGTCTGTGGTGGAGGCTCAGAAAGAAACAGCGTGTCAAGCCT		426
Oy	39	GLIUEVALIAIALASERIALARGALIAIALVALIYASPMETAEPLUGLINALAEUHLIS		58
Db	427	CCATTAAATACGACAAACCCCTTCTCTGCT-----		456
Oy	59	GLIARGIYVALIALAEUTHYRILEALHMETGLIAPSPGINDIYSERGLYSEURLETHR		78
Db	457	-----GCCATTGACACGTGCTGCTGCTTAATTAATCTTCAACCTCTTCTCTGCT		504
Oy	79	GLIYIARGTRYSERILEASPALEULEARGIYGLUTYRILEANSER-----		95
		:::	:::	
Db	505	CCACTACAAACCCGATCGACACCTCTTCAACCTCAATACATCAACCCTTGTTTAT		564
Oy	96	-----PROALVALARGTHRASPPIYRHYRTPROAGTYGLUTHRHPALAGU---		112
Db	565	GGCTTACCAAGTAGCTCTAGCGATGTGAATCTTCCATTGCTTCTTAGAGTTCCGCTTACA		624

```
Qy 113 ---ThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThr---LeuAsnAla 130
Db 625 ATCCAGCTCGGGGTTTGGATCTTCAGCTGAATTAATGCTTGGATTAGCTTTTCATCC 684
Qy 131 ProAlaLeuSerArgThrGlnSerAspGlySerGlySerArgSerSerLeuGlyLeuAsn 150
Db 685 GGGGTTTGTCCAAATGAAAGCTAGATGATTAATACGGTTACAGAACTGGTTGGTTCAAC 744
Qy 151 IleGlyGlyMetGlyAspTyrArgAsnGluThrLeuThrThrAsnProArgAspThrAla 170
Db 745 AATACGCTCTTTCAAGTTAATACATCCACCACTAGTACTACTACTCCAGCTATGCTTCT 804
Qy 171 PheLeu---SerHisLeuValGlnThrValPhePheLeuArgGlyIleAsp 186
Db 805 CTTTGACGCTCTTCTTGCTTCAACAGAGTTTCATGACAGATGGTGTGAT 855

RESULT 13
US-10-424-599-111877
; Sequence 111877, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OR INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111877
; LENGTH: 1154
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72033C.1
US-10-424-599-111877

Alignment Scores:
Pred. No.: 0.145 Length: 1154
Score: 99.50 Matches: 51
Percent Similarity: 37.63% Conservative: 19
Best Local Similarity: 27.42% Mismatches: 69
Query Match: 6.05% Indels: 47
DB: 18 Gaps: 8

US-10-617-835-4 (1-320) x US-10-424-599-111877 (1-1154)
Qy 23 ThrGlyIleProSerHisGlyGlyValysArgPheAlaValGluGlnGluLeuValAla 42
Db 343 ACTGGCAGGTCAAGAGGTTTGGCTTGTAACTTTTGCACAAAGTAGAGAT----- 393
Qy 43 AlaSerAlaArgAlaAlaValAllysAspMetAspLeuGlnAlaLeuHisGlyArgLysVal 62
Db 394 -----GCATCTCTGCATTCACAGGCGATGATGGCCAGATCTTCATGTCGAGAGATA 447
Qy 63 AlaLeuTyrIleAlaThrMet-----GlyAspGlnGlySerGly 75
Db 448 CGGGTGAAATTATGCTACAGAAAGTCACGTCACAGGTTTGTGTGTGATGTGATATAGGC 507
Qy 76 SerLeuThrGlyGlyArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSer 95
Db 508 AGTGTGCGACGTGGTGGCTACATATAGGGT-----GAAACTATGCAAGTGA 555
Qy 96 ProAlaValArgThrAspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSer 115
Db 556 GGTGATATACAGACGTAGTATGAGCTATATAGG-----GGTGAACATATGGAAGT 606
Qy 116 GlyGlyLeuThrGlyLeuThrThrSerLeuSerThrIleAsnAlaProAlaLeuSerArg 135
Db 607 GCGCGTTAT--AATGTTACTAGCAGCTATATAGTGTGCGCAATGCT-----GAA 651
```

```
Qy 136 ThrGlnSerAspGlySerGlySerArgSerSer-----LeuGlyLeuAsnIleGlyGly 153
Db 652 ACTAGTTAACTACTGGTGGTGAATGCTAGTATTAATCAATTCATGAAACTCTGCTGCA 711
Qy 154 MetGlyAspTyrArgAsnGluThrLeuThrThrAsnProArgAspThrAlaPheLeuSer 173
Db 712 GTTTTGGCTCAGCTACGCGTGAATTCACGCAACCAAAATGACACA----- 759
Qy 174 HisLeuValGlnThrValPhePheLeuArgGlyIleAspValAlaSerProAlaAsnAla 193
Db 760 -----GCAGGTGCA 768
Qy 194 AspThrAspValPheIle 199
Db 769 GACAATGATGAAATTCATT 786

RESULT 14
US-10-282-122A-19283
; Sequence 19283, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OR INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPITRA-034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19283
; LENGTH: 2622
; TYPE: DNA
; ORGANISM: Enterobacter cloacae
US-10-282-122A-19283

Alignment Scores:
Pred. No.: 0.588 Length: 2622
Score: 99.00 Matches: 64
Percent Similarity: 38.49% Conservative: 53
Best Local Similarity: 21.05% Mismatches: 147
Query Match: 6.02% Indels: 40
```

```

DB:      17          Gaps:      13
US-10-617-835-4 (1-320) x US-10-282-122A-19283 (1-2622)
QY      30 GlyGlyLysArgPheAlaValGluGlnIleuValAlaAlaSerAlaArg-----Ala 47
      151 GCGCGCAAGAGCTTACCCGCG-----ACGGTAGACGCCGGAAGCAAAATGACCGCC 201
QY      48 AlaValLysArgMetAspLeuGlnAlaLeuHisGlyArgLysValAlaLeuTyrIleAla 67
      202 ACGGTCGCCAGCGCGCATCTGCGGGTCTGAAGAAGCGGACGCGACGCTGCAGGTGAGAC 261
QY      68 ThrMetGlyAspGlnGlySerGlySerLeuThrGlyArgTyrSerIleAspAlaLeu 87
      262 GTCACCAACGTTAAAGCAACAGCCGCTCGCGGGCGCGGAGTACAGCTGATGCAAC 321
QY      88 IleArgGlyLysTyrIleAsnSerProAlaValArgThrAspTyrThrTyrProArgTyr 107
      322 GCGCGCGTCGTCACCATCAATATCTCGCG-----ACGACGATATCTGAAACGCGCC 375
QY      108 GluThrThrAlaGluThrThrSerGlyLysLeuThr-----GlyLeuThrThr 123
      376 GAAGGGAAGAGCATCTGACCGTATCCGCGACACCGCGGACCGGCGGCGACGCTGAGC 435
QY      124 SerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGlySer 143
      436 ACCGATGCGCTGAACGCGCAAGACTAACACACACCGCTGACGCGACGCGCGCTGAGC 495
QY      144 ArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGluThrLeuThr 163
      496 CTGAACGTCGCGCGCGCGCATCTGCGTGAATTACCGAC---GCGAGCGCTACCGCTTACC 552
QY      164 ThrAsnProArgAsp-----ThrAlaPheLeuSerHisLeuValGlnThrVal 179
      553 GCATCGCTGACGCAAGCGGCAACCGCGCGTGCCTGATCAACATCTG----- 603
QY      180 PhePheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIle 199
      604 -----GCGGTAGATATCACCGCGCGCTGCG-----GTGACCATC 636
QY      200 AsnIleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGlu 219
      637 AATACCGTGGCGGCGGACGATGTTATTAAATCCGACGAA-----CACGCGACGCGC 687
QY      220 ThrLeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLeu 239
      688 CAGATCATCATCGCGGACCGCCACCGCGCGGACGCGGAGTAAAGTGAACCTGCACCATC 747
QY      240 LeuIleLysProLysThrAsnAlaPheGlnAlaIleTyrLysGluAsnTyrAlaLeuTyr 259
      748 GCGGCGCAACCTTACACCAACCGCTGCTGATGCGCG-----GGGAACTGGAGCGTGGC 801
QY      260 MetGlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPhe 279
      802 GTA---CGGCGAGGCGTATTCGCGCCTC-----AGCGACGCGACCGTAAACCGTCACT 852
QY      280 SerAspIleArgProTyrGlyAsnHisGlyThrGlyAsnSerAlaProSerValGluAlaAsp 299
      853 GCGTCGCTCACGACGCGCGCGGCAACCGGTACCGGACGACATGACCGTGAAT 912
QY      300 AsnSerHisGluGlyTyrGlyTyrSer-----AspGlnAlaValAlaArgLysHis 315
      913 ACCGCTTCGCCATCGGTGCGCTTCAACGCGCATCAGCGATGACACGCTCTGAACGCGGTA 972
QY      316 ArgGlnGlyGln 319
      973 GAAAAAGGTCAAG 984
DB

```

```

RESULT 15
US-10-146-772-69/c
; Sequence 69, Application US/10146772
; Publication No. US20030124698A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Short, Jay
; APPLICANT: Weiner, David
; APPLICANT: Chaplin, Jennifer
; APPLICANT: Chi, Ellen
; APPLICANT: Milan, Aileen
; APPLICANT: Desantis, Grace
; APPLICANT: Madden, Mark
; APPLICANT: Burk, Mark
; TITLE OF INVENTION: Nitrlases
; FILE REFERENCE: Docket No. US20030124698A1 DIV-013US
; CURRENT APPLICATION NUMBER: US/10/146,772
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/309,006
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US 60/351,336
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/300,189
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 09/751,299
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/254,414
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/173,609
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 386
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample
; US-10-146-772-69

Alignment Scores:
Pred. No.: 0.141          Length: 939
Score: 98.50             Matches: 73
Percent Similarity: 35.20%      Conservative: 40
Best Local Similarity: 22.74%   Mismatches: 118
Query Match: 5.99%           Indels: 90
DB: 15                     Gaps: 11

US-10-617-835-4 (1-320) x US-10-146-772-69 (1-939)
QY      32 LysArgPheAlaValGluGlnIleuValAlaAlaSerAla-----Arg 46
      936 CCGCGATTCGCGCTCCACCGCGCGGACGCGCTGACGCGCATCGACGACCTCCGC 877
QY      47 AlaAlaValLysArgMetAspLeuGlnAlaLeuHisGlyArgLysValAlaLeuTyrIle 66
      876 GCGGCTGTAATGCGCGACGAGTCA----- 850
QY      67 AlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyArgTyrSerIleAspAla 86
      849 CCAGCGCTTGCGATGCAAGCGCGCTGCGGACGATCGGCGGACACGATCCCTCCG 790
QY      87 LeuIleArg-GlyGlyTyrIleAsnSerProAlaVal----- 98
      879 ATCGTAGAGCGCGCGCGGATTAATCTCGCCCGAGGTGAGCTGACGATCGCGACCGCC 730
QY      99 ArgThrAspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyLys 118
      729 GCGCGCAACACCTCTTCCCGCGGGGTAAGCTGCACGCGGGAATGTCGCGGGAACCGGA 670
QY      118 uThrGlyLeuThrThrSerLeuSerThrLeuAsnAla-ProAlaLeuSerArgThrGlns 138
      669 CGCGCGATGAACCTCGCGACCGACGACGACGACGCGCGCGCTGATCGCGATGT---- 614
QY      138 eArgGlySerGlySerArgSerSer-LeuGlyLeuAsnIleGlyGlyMetGlyAspTyr 157
      613 -----GCGCATGCTCGGAGCGACGCGCTCGGAGTC---ATCGAC----- 577
QY      158 ArgAsnGluThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGln 177

```

```
Db 576 -----CGTCGGCGCCAGATCTGCGGTCCACCCCTGGTAGAC 538
Qy 178 ThrValPhePheLeuArgGlyIleAspValSerProAlaSerAlaAspThrAspVal 197
Db 537 GCGGTA-----GCGCGGAGGGGATTGCTTCCAGCAGATCAGTCCCGCAT--- 487
Qy 198 PheIleAsnIleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsn 217
Db 486 -----CCGTCCCGGTGCGTCTCCACAC 463
Qy 218 AlaGluThrIleLeuLysAlaGluThrIleLeuGluTyrPheAlaValAspArgThrAsnLys 237
Db 462 GCCGAGGTCTTGACCGTCGCGATCCATGGAAACGCCGCTGCTGTGTCGTCGCGCATGAG 403
Qy 238 -----LysLeuLeuLys 242
Db 402 CTGGCGGTGCGGTGCGAGAGCCCGAGGGGTCAAGCATCGTGTGTAAAGCGA 343
Qy 243 ProLysThrAsnAlaPheGluAlaAlaTyrIleAsnTyrAlaLeuTrpMetGlyPro 262
Db 342 CCGCGGCTTTCGCTTTCGCGCTGTCACGCGCATCAC-----GCA 301
Qy 263 TyrIleValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSerAspIle 282
Db 300 TACCACGTCATGCTGCGGACGATCGACGACGCGGTGACACAGCGGCGCGGACGCTC 241
Qy 283 ArgProTyrGlyAsnHisThr-----GlyAsn 291
Db 240 GAGCGAGCTGGCCACATCCGCTCCAGAGCTCGTCAAGCGCGCAATCCGGCGGCGCC 181
Qy 292 SerAlaProSerValGluAlaAspAsnSerHisGluGlyTyrGlyTyrSerAspGlu 310
Db 180 TCGGCGCCACCGCTTCGACGGGTAGAGGACACGAAAGGCTCGGGCAGCAGACGAG 124
```

Search completed: August 19, 2005, 02:32:36  
Job time : 5097 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 18, 2005, 23:14:35 ; Search time 198 Seconds  
(without alignment)  
2644.488 Million cell updates/sec

Title: US-10-617-835-4  
Sequence: 1 MRARLLPILBSVFLSACG.....SHEGYGSDAVRQROGCP 320

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
MODE=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool.p/US1061735/rnat.18082005.115609.8386/app.query.fasta\_1.519  
-DB=Issued Patents NA -QFMT=fasta -SUPRT=p2n.rnt -MINMATCH=0.1 -LOOPCL=0  
-IOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1061735 @CGN 1 1.105 @rnat.18082005.115609.8386 -NCPD=6 -ICPD=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGCLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/ECTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1645	100.0	1136	US-09-043-302-6	Sequence 6, Appli
2	1645	100.0	3287	US-09-043-302-1	Sequence 1, Appli
3	106.5	6.5	573	US-09-543-681A-268	Sequence 268, Ap
4	106.5	6.5	1521	US-09-489-039A-5599	Sequence 5599, Ap
5	101	6.1	1719	US-09-248-796A-1633	Sequence 1633, Ap
6	98.5	6.0	7109	US-09-902-540-922	Sequence 922, App
7	97.5	5.9	2196	US-09-252-991A-5953	Sequence 5953, Ap
8	97.5	5.9	3036	US-09-252-991A-5934	Sequence 5934, Ap
9	97	5.9	1548	US-09-489-039A-3464	Sequence 3464, Ap
10	97	5.9	4063	US-09-902-540-595	Sequence 595, App
11	95	5.8	1896	US-08-426-630-33	Sequence 33, Appl
12	95	5.8	4748	US-08-426-630-29	Sequence 29, Appl

C 13	95	5.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
14	94.5	5.7	1089	3	US-09-071-035-157	Sequence 157, Appl
15	93.5	5.7	1491	4	US-09-252-991A-1243	Sequence 1243, Ap
C 16	93.5	5.7	1551	4	US-09-252-991A-1362	Sequence 1362, Ap
17	93	5.7	2346	4	US-09-252-991A-6782	Sequence 6782, Ap
18	93	5.7	3619	3	US-08-377-503-1	Sequence 1, Appli
19	93	5.7	3619	3	US-08-178-019-1	Sequence 1, Appli
20	93	5.7	7152	4	US-09-492-709A-142	Sequence 142, App
21	92.5	5.6	2780	4	US-09-841-786-10	Sequence 10, Appl
22	92.5	5.6	9726	4	US-09-841-786-8	Sequence 8, Appli
23	92.5	5.6	11130	4	US-09-841-786-15	Sequence 15, Appli
24	92	5.6	3484	4	US-09-949-016-1893	Sequence 1893, Ap
C 25	91.5	5.6	1974	4	US-09-489-039A-6189	Sequence 6189, Ap
26	91.5	5.6	2201	4	US-09-489-039A-5812	Sequence 5812, Ap
27	91	5.5	3600	5	PCT-US95-13749-1	Sequence 1, Appli
28	91	5.5	15664	1	US-08-402-282-3	Sequence 3, Appli
29	91	5.5	15664	1	US-08-508-004-3	Sequence 3, Appli
30	91	5.5	15664	1	US-08-402-066-3	Sequence 3, Appli
31	91	5.5	15664	1	US-08-402-066-3	Sequence 3, Appli
32	90	5.5	4308	3	US-08-938-291A-8	Sequence 8, Appli
33	90	5.5	4308	4	US-09-589-619-8	Sequence 8, Appli
34	90	5.5	1830121	4	US-09-557-884-1	Sequence 1, Appli
35	90	5.5	1830121	4	US-09-643-990A-1	Sequence 1, Appli
36	89	5.4	3687	4	US-09-463-402-1	Sequence 1, Appli
37	89	5.4	3687	4	US-09-889-572-1	Sequence 1, Appli
38	89	5.4	3687	4	US-09-117-447-1	Sequence 1, Appli
39	89	5.4	3768	4	US-09-463-402-12	Sequence 12, Appl
40	89	5.4	4079	4	US-09-016-434-1412	Sequence 1412, Ap
41	89	5.4	4988	4	US-09-463-402-11	Sequence 11, Appli
42	89	5.4	6855	4	US-09-252-991A-6976	Sequence 6976, Ap
C 43	89	5.4	26659	4	US-09-902-540-1237	Sequence 1237, Ap
44	88.5	5.4	1134	4	US-09-540-236-1714	Sequence 1714, Ap
45	88.5	5.4	2643	4	US-09-486-072-6	Sequence 6, Appli

## ALIGNMENTS

RESULT 1  
US-09-043-302-6  
Sequence 6, Application US/09043302  
Patent No. 6617128  
GENERAL INFORMATION:  
APPLICANT: MEYER, Thomas F.  
APPLICANT: RUDEH, Thomas  
APPLICANT: SCHUBERPLUG, Ina  
APPLICANT: MAIER, Jurgen  
APPLICANT: EICKERNJAGER, Sandra  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impact  
TITLE OF INVENTION: the Adhesion of Neisseria Cells to Human Cells  
FILE REFERENCE: 0147-172P  
CURRENT APPLICATION NUMBER: US/09/043.302  
CURRENT FILING DATE: 1998-06-08  
EARLIER APPLICATION NUMBER: PCT/EP96/04092  
EARLIER FILING DATE: 1995-09-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 1136  
TYPE: DNA  
ORGANISM: Neisseria gonorrhoeae  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (135)..(1094)  
US-09-043-302-6  
Alignment Scores:  
Pred. No.: 1.63e-185  
Score: 1645.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Length: 1136  
Matches: 320  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-617-835-4 (1-320) x US-09-043-302-6 (1-1136)

QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 Db 135 ATGGGGGACGGCTGCTGATACCTATCTCTTTTTCGGTTTATTTATTTATCCGCCCGGG 194  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyGlyLysAspPheAlaValGlnGlnGlnLeu 40  
 Db 195 ACACGACAGGATTCATTCGATGCGGAGGCAACGCTTCGCGGTGACACAACTT 254  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 Db 255 GTGGCCGCTTCGCGAGAGTCCGCTTAAAGCATGATTTTACAGGACATTCACGACGA 314  
 QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
 Db 315 AAAGTTGATGTATGCAATTGCAACTATGGGCGACCAAGGTTTCAGGCAGTTTGACAGGGGT 374  
 QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
 Db 375 CGCTACTCATTTGATGACATGATTCGCGGCAATTCATTAACAGCCCTGCGCTCGCAC 434  
 QY 101 AspTyrThrTyrProArgTyrGlyThrThrAlaGlnThrThrSerGlyLysLeuThrGly 120  
 Db 435 GATTACACCTATCCGCGTTACGAACCAACCGCTGAACCAATCAGCGGTTTGACGGGT 494  
 QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 Db 495 TTAAACCACTTCTTATCTTACACTTATATGCCCCCTGACCTCGCGCAACCAATCAGACGGT 554  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGly 160  
 Db 555 AGCGGAATAGAGAGAGCTGGGCTTAATATTTGGCGGAGATGGGGATTTATCGAAATGAA 614  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
 Db 615 ACCTTGAAGACCAACCCCGCGACACTGCTTCTTCCCACTGGTACAGACCGCTATTT 674  
 QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
 Db 675 TTCGCGCGGATAGACGTTGTTCTCTCGCAATGCGCATACAGATGTTTATTAAAC 734  
 QY 201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHisLeuTyrAsnAlaGlyThr 220  
 Db 735 ATCGACGATTCGGAAGCATATGCAACGAACCGAAATGCACTTATACAAATGCCAAACA 794  
 QY 221 LeuLysAlaGlnThrLysLeuGlyLysPheAlaValAspArgThrAsnLysLeuLeu 240  
 Db 795 CTGAAGGCCCAAACTGGAATATTTCGACATGACAGAACCAATMAAAATTGCTC 854  
 QY 241 IleLysProLysThrAsnAlaPheGlnAlaAlaTyrLysGlyLysAsnTyrAlaLeuTrpMet 260  
 Db 855 ATCAAAACCAAAACCAATGCGTTGAGCTGCTTAAAGAAATTTAGCGCATTTGGATG 914  
 QY 261 GlyProTyrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer 280  
 Db 915 GGGCGGTATTAAGTAGCAAAAGAAATCAACCGACGGAAGATTATAGTGATTTCTCC 974  
 QY 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGlnAlaAspAsn 300  
 Db 975 GATATCCGGCATACGCGCAATCATACGGGTAACTCCGCCCATCGTGAAGCGTATTAAC 1034  
 QY 301 SerHisGlyGlyTyrGlyTyrSerAspGlnAlaValArgGlnHisArgGlnGlyPro 320  
 Db 1035 AGTCATGAGGGGATGATGATACAGCATGACGATGCGACAACTATGACAAAGGGCACTT 1094

RESULT 2

US-09-043-302-1

; Sequence 1, Application US/09043302

; Patent No. 6617128

; GENERAL INFORMATION:

; APPLICANT: MEYER, Thomas F.

; APPLICANT: RUDEL, Thomas  
 ; APPLICANT: SCHUEPFLING, Ina  
 ; APPLICANT: MAIER, Jurgen  
 ; APPLICANT: EICKENBAGER, Sandra  
 ; APPLICANT: SCHWAN, Thomas  
 ; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
 ; TITLE OF INVENTION: the Adhesion of Neisseria Cells to Human Cells  
 ; FILE REFERENCE: 0147-172P  
 ; CURRENT APPLICATION NUMBER: US/09/043,302  
 ; EARLIER APPLICATION NUMBER: PCT/EP96/04092  
 ; EARLIER FILING DATE: 1998-06-08  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3287  
 ; TYPE: DNA  
 ; ORGANISM: Neisseria gonorrhoeae  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (136)..(447)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (583)..(1542)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1585)..(3111)  
 ; US-09-043-302-1  
 Alignment Scores:  
 Pred. No.: 9,45e-185 Length: 3287  
 Score: 1645.00 Matches: 320  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-617-835-4 (1-320) x US-09-043-302-1 (1-3287)  
 QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
 Db 583 ATGGGGGACGGCTGCTGATACCTATCTTCTTTTTCGGTTTATTTATTTATCCGCTCGCGG 642  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyGlyLysAspPheAlaValGlnGlnGlnLeu 40  
 Db 643 ACACGACAGGATTCATTCGATGCGGAGGCAACGCTTCGCGGTGCAACAAAGAACTT 702  
 QY 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 Db 703 GTGGCCGCTTCGCGAGAGCTGCGCTTAAAGACATGATTTTACAGGCATTTACAGGACGA 762  
 QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
 Db 763 AAAGTTGATGTATGCAATTGCAACTATGGGCGACCAAGGTTGAGCAAGTTCAGAGGGGT 822  
 QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
 Db 823 CGCTACTCATTTGATGACATGATTCGCGGCAATTCATTAACAGCCCTGCGCTCGCAC 882  
 QY 101 AspTyrThrTyrProArgTyrGlyThrThrAlaGlnThrThrSerGlyGlyLeuThrGly 120  
 Db 883 GATTACACCTATCCGCGTTGCAACCAACCGCTGAACCAATCAGCGGTTTGACAGGGT 942  
 QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
 Db 943 TTAAACCACTTCTTATCTTACACTTATATGCCCCCTGCACTTCGCGCAACCATGACAGGT 1002  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGly 160  
 Db 1003 AGCGGAAGTAGAGACACTGCGGCTTAATATTGGCGGATGGGGGATTTATCGAAATGAA 1062  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180

Db 1063 ACCTTGACGACCAACCGCGCGACACTGCTTTCTTCCCACTGTGTACAGACCGGATTT 1122  
 QY PheLeuAArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValIleIleAsn 200  
 Db 1123 TTCCGCGCGGATAGACTGTCTTCTCTGCCAATGCCGATACAGATGTGTTTATAC 1182  
 QY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisIleuTyArgAlaGluThr 220  
 Db 1183 ATCGACGATTTGCGAACGATACGACAGAACCGAATACCTTATACATCCGAAACA 1242  
 QY 221 LeuValAlaGlnThrIleLeuGluIuTyPheAlaValAspArgThrAsnIleIleuLeu 240  
 Db 1243 CTGAAGACCCCAACAACTGGAATATTTCCGAGTAGACAGAACCAATTAATAATGCTC 1302  
 QY 241 IleLeuProIleThrAsnAlaPheGluAlaAlaIleTyArgGluAsnTyArgIleuThrMet 260  
 Db 1303 ATCAAAACCCAAACCAATGCCGTTGAAAGCTCTTAAGAAATTTACGCAATTTGGATG 1362  
 QY 261 GlyProTyArgValSerIleGlyIleIleuProThrGluGlyLeuMetValAspPheSer 280  
 Db 1363 GGGCGGATTAAGTAAGCAAGAAATCAAAACCGAAGGATTAATGTCGATTTCTCC 1422  
 QY 281 AspIleArgProTyArgIleAsnIleThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
 Db 1423 GATATCCGCGCATACGGCAATCATACGGGTACTCCGCCCATCCGTAGAGCTGATAC 1482  
 QY 301 SerHisGluGlyTyArgIleTySerAspGluAlaValArgGlnHisArgGlnGlyGlnPro 320  
 Db 1483 AGTCATGAGGGGTATGATGATACGACGATGAAAGTCGCGCAACATATGACAAAGGACACT 1542

## RESULT 3

US-09-543-268  
 ; Sequence 268, Application US/09543681A  
 ; Patent No. 6605709  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GARY BRETON  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.1002-001  
 ; CURRENT APPLICATION NUMBER: US/09/543,681A  
 ; CURRENT FILING DATE: 2000-04-05  
 ; PRIOR APPLICATION NUMBER: US 60/128,706  
 ; PRIOR FILING DATE: 1999-04-09  
 ; NUMBER OF SEQ ID NOS: 8344  
 ; SEQ ID NO 268  
 ; LENGTH: 573  
 ; TYPE: DNA  
 ; ORGANISM: Proteus mirabilis  
 US-09-543-681A-268

## Alignment Scores:

Pred. No.: 0.00476 Length: 573  
 Score: 106.50 Matches: 50  
 Percent Similarity: 40.58% Conservative: 34  
 Best Local Similarity: 24.15% Mismatches: 84  
 Query Match: 6.47% Indels: 39  
 DB: 4 Gaps: 9

US-10-617-835-4 (1-320) x US-09-543-681A-268 (1-573)

QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleuSerAlaCysGly 20  
 Db 16 CTTAAATAAAGATGTAGTTAAAGCTCTATTACAGTATCTGTTTATACCTGAGCGGT 75  
 QY 21 ThrLeuThrGlyIleProSerHisGlyGlyTyArgPheAlaValGlnGlnIleu 40  
 Db 76 GTTGCCTTCGACGACGACTATTGTTAATGTTGTA--ATTAAATTTACTGAGAAAT 132  
 QY 41 ValAlaIleSerAlaArgAlaValIleAspMetAspLeuGlnAlaLeuHisGlyArg 60  
 Db 133 GTTAATGCTGACGTGTGTGTGTGATCAAAATCAATATGATCAAACTGTTAATTTGGGCA 192  
 QY 61 LysValAlaLeuTyrlleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80

Db 193 TATGCTACGCAATTTGATAGTGTAGGAA-----ACAGTGTGT 234  
 QY 81 ArgTySerIleAspAlaLeuIleArgGlyGlyTyrlleAsnSerProAlaValArgThr 100  
 Db 235 AATACTGATTTT-----TATATCAATTTAGAAAGCTTGATACG 273  
 QY 101 AspTyThrTyProArgTyArgIleThrAlaGluThrThrSerGlyGlyLeuThrGly 120  
 Db 274 -----ACTGCGACGAATGCTTCGTTCAATTTTCAGGT 309  
 QY 121 LeuThrThrSerLeuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspIle 140  
 Db 310 GTAAGTATTTCA--AATGATAAACAGTATTAAGCTGTATACCAATATTAACAAGCTGT 366  
 QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyIleMetGlyAspTyArgAsnGlu 160  
 Db 367 GCGGCTGCTGCGACGCGCGTGTGTAATAAT-----ACAGACCATACAGGAA 417  
 QY 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisIleuValGlnThrVal--- 179  
 Db 418 GTGTTA-----CCACAGACGCGCTCTGTTTCTACACGAAACAGTTAATGAT 468  
 QY 180 -----PhePheLeuArg-----GlyIleAspValValSerPro 190  
 Db 469 GATCAATAATACACTTAATTGTTGTCACGCTTATAATCTACATTAGATACGTTACCCCT 528  
 QY 191 AlaAsnAlaAspThrAspVal 197  
 Db 529 GGTCACTGCTGATGTGATGTA 549

## RESULT 4

US-09-489-039A-5599  
 ; Sequence 5599, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; CURRENT FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 5599  
 ; LENGTH: 1521  
 ; TYPE: DNA  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-5599

## Alignment Scores:

Pred. No.: 0.0239 Length: 1521  
 Score: 106.50 Matches: 60  
 Percent Similarity: 37.64% Conservative: 39  
 Best Local Similarity: 22.81% Mismatches: 85  
 Query Match: 6.47% Indels: 79  
 DB: 4 Gaps: 14

US-10-617-835-4 (1-320) x US-09-489-039A-5599 (1-1521)

QY 9 IleLeuPheSerValPheIleuSerAlaCysGlyThr-----Leu 22  
 Db 769 GTATTAACTGCTGTGTTATTACATCTCTGTGTAACTGGTTTTCATACAGAAAGATTG 828  
 QY 23 ThrGlyIleProSer---HisGlyGlyGly---LysArgPheAlaValGlnGlnIleu 40  
 Db 829 AATGGGAACGATGATTCATTCATGCGACAGAAATGAAGGTGCTTTCAACAGCA--- 885  
 QY 41 ValAlaIleSerAlaArgAlaValIleAspMetAspLeuGlnAlaLeuHis---Gly 59  
 Db 886 -----AAAAAAGATTAATATGTCCTCATCATACATTGCT 921

```

QY 60 arglysvAlaIaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGly 79
Db 922 GAGAAAGTCAGCGATATATATATATGCGCAATG----- 951
QY 80 GlyArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArg 99
Db 952 -----ACTATTCATGTCGTTTAGAGATATACACTACATCTCCATATATACGA 996
QY 100 ThrAspTyrThrTyrProArgTyrGluThrThrAlaGluThrThr----- 114
Db 997 CACGAC--GGGTGCCCAAGATCAAAATGTATGTGTATATACCGTTGAGAAAAAGCCT 1053
QY 115 -----SerGlyLeuThrGly 120
Db 1054 AAAAATATATAGTAAGCTCTGTAGATGATATCCGTTATACATTTTAGGACGATATCATGT 1113
QY 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db 1114 TTGACCGAAAAATCTGTCAACAACATACTTCTTGACAGCAGATTTCACAGCTAATCT 1173
QY 141 SerGlySer-----ArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAsp 156
Db 1174 GATGCGCTCCGTCATATTTATACGAAACCAAGCGTCATTATCT--GGCGGGTTCGATAGC 1230
QY 157 TyrArgAsnGluThrLeuThrThrAsnProArgAspThrAlaPheLeuSer---HisLeu 175
Db 1231 GAAAAAGAGCGCGGATTCATCTTAATGCGGGGTGAAGTGGCTTCATGATGATTACTGTT 1290
QY 176 ValGlnThrValPhePheLeuArgGlyIleAspValSerProAlaAsnAlaAspThr 195
Db 1291 ATTCTTAATGCAAGTTGTCAGTAGGGGCTCG----- 1320
QY 196 AspValPheIleAsnIleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeu 215
Db 1321 ---ATTTTATATAGCTGTATGATCAAAATGTAATGCGCAA-----ATA 1359
QY 216 TyrAsnAlaGluThrLeuLeuPheIleGlnThrIlySLeuGlu-----TyrPheAlaVal 232
Db 1360 ACTGACACAGCGCATGCTCTCAAGTAGAGATTAATTAAGACAGATTAATATTTGTTTTT 1419
QY 233 AspArgThr 235
Db 1420 GACAGGTCA 1428

RESULT 5
US-09-248-796A-1633
; Sequence 1633, Application US/09248796A
; Patent No. 6747237
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1633
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1633

Alignment Scores:
Pred. No.: 0.132 Length: 1719
Score: 101.00 Matches: 65
Percent Similarity: 35.16% Conservative: 31
Best Local Similarity: 23.81% Mismatches: 109
Query Match: 6.14% Indels: 68
DB: 4 Gaps: 12

```

US-10-617-833-4 (1-320) x US-09-248-796A-1633 (1-1719)

QY 64 LeuYrjlllealAthrmetGlyAspGlnGlySerGlySerLeuThrGlyGlyArgTySer  
Db 157 CTAAGTACTACTACTACTATAGCATTAATAATTATACATCTAATAGATCCGGTACAGTTAAA  
QY 84 lLeasrAlaLeuLeaArgGlyGlyUryrllleasnsrProAlaValArgThrAspTyThr  
Db 217 ATTGATTTGT-----TCTGGAGCATTTATTTGCCAATGATGATTTTATTCGAGT--ACT  
QY 104 TyzProArgTyGlyUthrThralaGluThrThrSerGlyGlyLeuThrGlyLeuThrThr  
Db 268 GCACCT-----GAATTAATCTAGTAATAATTACAAAT-----TCC  
QY 124 SerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSer-----Asp  
Db 301 TCTTCTTCAACTTCGCGGACACCTCTCTGTCTAAATATACATGCTGGATCTTGAAAGATGAC  
QY 140 GlySerGlySerArgSerSerLeuGlyLeuAsnllleglGlyMetGlyAspTyArgAsn  
Db 361 ACAGAGAAATCTGAATTGACACAGAGAAATGGTGAGGTTTAAAGTAT-----  
QY 160 GlUthrThrThrThrAsnProArgAspThrAlaPheLeuSerThrSleuValGlnThrVal  
Db 412 TCTTCAAGCACCAACGGTGGAGAATGACGAAATCAATTTACAACAATTCGACAAACACT  
QY 180 PhePhe-----LeuArgGlylLeaSpValValSerProAlaAsnAlaAspThrAsp  
Db 472 TTTATATGAGGTGCTCTATCCAAACAAATATACCTTACCAACAAACATGATGAATC  
QY 197 ValPheIleAsnLeaSpValPheGlyThrIleArgAsnArgThrGlnMetThrSleuTy  
Db 532 ATATATTTATCTGTGGATTAAGGTGTCATTA-----  
QY 217 AsnAlaGluThrLeuLysAlaGlnThrIrySleuGluIryrPheAlaValaAspArgThrAsn  
Db 565 -----ACCAATTAAGTTTATTTACTTAATCTTCATCAAT  
QY 237 LysAlaLeuLeuLeuLysProLysThrAsnAla-----  
Db 601 AGGAAAAATCTTCTTGAAGAAAACCACTGGTGTGGAGGTGGGGTAATTACATTTCCAG  
QY 248 ---PheGluAlaAlaTyrrGluAsnTyrrAlaLeuTrpMetGlyProTyrrLysValSer  
Db 661 GATTTCCAACGGAATATTTTAAATATGAAGTATGAAGTATCATGATTAATTAATACAA  
QY 267 LysGlylLeaSerProGlnGlyLeuMetValaAspPheSerAspIleArgPro-----  
Db 721 AAATCAATTAATAAA-----TCTAATATTAATTCCTCTATA  
QY 285 ---TyrrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerHisGlu  
Db 757 ACGTAATTAATGGCAACTGAGAGAGGT-----CATAAA  
QY 304 GlTyrrGlyTyzSerAspGluAlaValaArgGlnHisArg 316  
Db 793 TTTTATTAATTTATGAATAAAACATTTAATAAACATAAA 831  
RESULT 6  
US-09-902-540-922  
; Sequence 922, Application US/09902540  
; Patent No. 6633447  
GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; FILE REFERENCE: 38-10 (15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883







[illegible]

```

Db      1168 GCGGCAGAAAGCTTTGGCAGACAGCCSACSTGCGGCATTTGAATAAAGTATAGGTAC 1227
        |||||                                     :|||
QY      308 SerArgGluAlaValArgGlnHisArgGlnGlyLeu 319
        ::|||                                     :|||
Db      1228 AACCTGTAAAGCAACAGCACAGCATGTATGAAGTGAA 1263
        ::|||                                     :|||

RESULT 10
US-09-902-540-595/c
; Sequence 595, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wisgand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIORITY FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 595
; LENGTH: 4063
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-595

Alignment Scores:
Pred. No.:          1.63              Length:       4063
Score:             97.00              Matches:       78
Percent Similarity: 35.36%            Conservative:   50
Best Local Similarity: 21.55%         Mismatches:    120
Query Match:       5.90%               Indels:        114
DB:                4                  Gaps:          15
US-10-617-835-4 (1-320) x US-09-902-540-595 (1-4063)
QY      1 MetArgLa--ArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCys 19
        |||||                               |||:::|||
Db      2940 ATGGCGGCCATCAAGCGCTGTCCTCCATCTC-----GCGTGGTGGTACCAGCATGT 2887
        |||||                               |||:::|||
QY      20 GlyThrLeuThrGlyIleProSerHisGlyGlyGlyLeuValArgPheAlaValGlnGln 39
        ::|||                               ::|||
Db      2886 TCTGACTCCGCGCTCCCCCGCATGACAGACTGACCGGAGTGGGCATGCATCACCAAGAG 2827
        ::|||                               ::|||
QY      40 LeuValAlaAlaSerAlaArgAlaAlaValLeuValMetLysMetLysGlnAlaLeuHisGly 59
        ::|||                               ::|||
Db      2826 GCGCTGCGCGCGCATCGTGGCGGAGAGCTATGCAAGTCTCCACCGCTTGGAGAGC 2767
        ::|||                               ::|||
QY      60 Arg-----LysValAla-----LeuTyrlleAla 67
        ::|||                               ::|||
Db      2766 ACGGCTGCGATTGCTGGCGCGTGGCGGAGGCGGCGCATGCGCCACCGGACTGAGCTGCC 2707
        ::|||                               ::|||
QY      68 ThrMetGlyMetGlnGlySerGlySerLeuThrGlyGlyArgTyrlSerlLeuAlaLeu 87
        ::|||                               ::|||
Db      2706 ACCCTGGGCGCAATACACCGGAGCGCCCAATCGGAGGACCTTCATCTTTCAGCGGAG 2647
        ::|||                               ::|||
QY      88 IleArg-----LysValAla-----LysValAla 89
        ::|||                               ::|||
Db      2646 GTCCGTGATTCGCCGCGAGCGGACTGTTGCGGCGCATGTCGAGTCTGAGTGTCTAACGCC 2587
        ::|||                               ::|||
QY      90 GlyLeuTyrlLysSerProAlaValArgThzArgTyrlTyrlProArgTyrlGluThr 109
        ::|||                               ::|||
Db      2586 GCGCTCCAGTGGCGGTATCCAGCTGCCGAGTSCATACGTCGGCGGCACTACGCGGCG 2527
        ::|||                               ::|||
QY      110 ThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThrThrSerLeuSerThrLeuValn 129
        ::|||                               ::|||
Db      2526 CAATCTGAT--GCTTGGGCGGACAGACCGGTACACCTGTAAGCTTCCGCGGAGGAGCAC 2470
        ::|||                               ::|||
QY      130 AlaProAlaLeuSerArgTyrlHnGlnSerArgGlySerGlySerArgSerLeuGlyLeu 149
        |||||                               :|||

```

Db 2469 CTCGGTGGGGGCTGCGCTCGATACAGAGGG----- 2437  
Gy 150 AsnIleGlyGlyMetGlyAspTyrArgAsnGluThrLeuThrThrAsnProArgAspThr 169  
Db 2436 CACATCTCCGGGTT----- 2422  
Gy 170 AlaPheLeuSerHisLeuValGlnThrValPhePheLeuArgGlyIleAspValValSer 189  
Db 2422 ----- 2422  
Gy 190 ProAlaAsnAlaAspThrAspValPheIleAsnIleAspValPheGlyThrIleArgAsn 209  
Db 2421 CCGCGCACCTCCGACAC-----TTCTCCGACCGCTC----- 2389  
Gy 210 ArgThrGluMetHisLeuTyrAsnAlaGluThrLeuValAlaGlnThrIleGluGlu--- 228  
Db 2388 -----CAGTGAAGAGCGCACCGGCGTTTGAAGCAACGTTCTCTCGCTC 2341  
Gy 229 -----TyrPheAlaValAspArgThrAsnValLeuLeuIleLys 242  
Db 2340 TCCACTTCACCGGCGCCTATCTCGCAGCGACAGCTC----- 2302  
Gy 243 ProLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTrpMetGlyPro 262  
Db 2301 CCTCCCGCGCTCGTGGAGCGGCTTCTCCACGAGCTTCACGCGCTTGAAGGCGACCA 2242  
Gy 263 -----TyrLeuValSerLysGlyIleLysProThrGluGlyLeuMetValAspPhe 279  
Db 2241 CCGCTCACGTTCCGATACCTCAGGCTTGTCCCGACG--GGCTTCACACTGATGCC 2185  
Gy 280 SerAspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAsp 299  
Db 2184 AGCGGTGCCGAGGAGACCCGACGACGAGCGCGCTCTTACGCTGCAATTCAG 2125  
Gy 300 AsnSerHisGluGlyTyrGlyTyrSerAspGluAla--ValArgGlnHisArgGlnGly 319  
Db 2124 GACGCGAAT-----GGCGGTGCTGTCGCGCACGTTCAACCTCAGCTCCAGGAGC 2074  
Gy 319 In 319  
Db 2073 GG 2072

RESULT 11  
US-08-426-630-33/c  
; Sequence 33, Application US/08426630  
; Patent No. 6656709  
; GENERAL INFORMATION:  
; APPLICANT: BLANCHE, FRANCIS, CAMERON, BEATRICE, CROUZET,  
; APPLICANT: JOEL, DEBUSCHE, LAURENT, LEVY SCHIL, SOPHIE,  
; APPLICANT: THIHAUT, DENIS  
; TITLE OF INVENTION: POLYPEPTIDES INVOLVED IN THE  
; TITLE OF INVENTION: BIOSYNTHESIS OF COBAMIDINS AND/OR COBAMIDES, DNA SEQUENCES  
; TITLE OF INVENTION: CODING FOR THESE POLYPEPTIDES, PREPARATION METHOD AND THEIR  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 555 13TH STREET, N.W.  
; CITY: WASHINGTON  
; STATE: DISTRICT OF COLUMBIA  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/426,630  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/916,151

; FILING DATE: 14-SEP-1992  
; APPLICATION NUMBER: PCT/FR91/00054  
; FILING DATE: 30-JAN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: F. P. CALVERT  
; REGISTRATION NUMBER: 28,557  
; REFERENCE/DOCKET NUMBER: 1290-7213  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 857-7887  
; TELEFAX: (202) 857-7929  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1896 base pairs  
; TYPE: Nucleic acid  
; STRANDEDNESS: Double  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Pseudomonas dentrificans  
; STRAIN:  
; INDIVIDUAL ISOLATE:  
; DEVELOPMENTAL STAGE:  
; HAPLOTYPE:  
; TISSUE TYPE:  
; CELL TYPE:  
; CELL LINE:  
; ORGANELLE:  
; FEATURE:  
; NAME/KEY: cobT  
; LOCATION: 2616-4511 bp of SEQ ID NO: 29  
; IDENTIFICATION METHOD:  
; OTHER INFORMATION:  
US-08-426-630-33

Alignment Scores:  
Pred. No.: 0.801 Length: 1896  
Score: 95.00 Matches: 73  
Percent Similarity: 31.10% Conservative: 34  
Best Local Similarity: 21.22% Mismatches: 99  
Query Match: 5.78% Indels: 138  
DB: 4 Gaps: 17

US-10-617-835-4 (1-320) x US-08-426-630-33 (1-1896)

Gy 14 PheIleuSerAlaCySGlyThrLeuThrGlyIleProSerHisGlyGlyLysArg 33  
Db 1854 TTCTGCTCGAAGAGTGGCGGCGACGTGTCGT-----CATCGCGCGGCAAGCTC 1804  
Gy 34 PheAlaValGluGln---GluLeuValAlaAlaSerAlaArgAlaAlaValLysAspMet 52  
Db 1803 ATCGGATCGACGATGAGTGAAGCGACGCGCATGAGCGGTACGTCGCGCGATACC 1744  
Gy 53 AspLeuGlnAlaLeuHisGlyArg-----LysVal 62  
Db 1743 GATCGCGACGAGTTCACCGCGGAGCGGTTTCATCTGCTGATGACCGCGCGAGGTG 1684  
Gy 63 AlaLeuTyrIleAlaIleThrMetGlyAspGlnGlySerGlySerLeuThrGlyIleArgTyr 82  
Db 1683 ACGCTCCAGATAGTTTCTGATTCGACGA-----CAGCGTGA 1645  
Gy 83 SerIleAsp--AlaLeuIleArgGlyGluTyrIleAsnSerProAla----- 97  
Db 1644 GTCTGACCGCGCGCGCGCTCCGAATCATCATCATGATGCGCGCGCTTTCGCGCGCGC 1585  
Gy 98 -----ValArgThrAspTyrThr 103  
Db 1584 CATCAGCGCTCATGCGCGCAATCAACGCTCGCGCTGATGTTTCTTTCGACAGGCC 1525  
Gy 104 TyrProArgTyrGluThr----- 109  
Db 1524 TTCCCGCATCATCAGCGCGAGATTGCGTGTGCGCGCGCGCAGGAGCGTGGCAACTT 1465

```

QY 110 -----ThralaGluThrThrseryL----- 116
DB 1464 GTAGACATGTGTCCAGAGTCCTTAGAGCGACCCGGGCTGTGTGCTCCGCCGCCAG 1405
QY 117 -----GlyLeuThrGlyLeuThrThrseryL 126
DB 1404 CCACTTCTGGGTACTGCGCCACCTTCCAGCCTTGGGTGTAAGAAACGAGATCTCAGC 1345
QY 127 ThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAsp-----GlySeryLys 143
DB 1344 ---CTTGAAGCGCGCGCTCGAGCGTGGCGCCGAGATATCGCGCGAGTGGCGCGAAC 1288
QY 144 ArgSerSerLeuGlyLeuAsnAlaLeuGlyMetGlyAsp---TyrArgAsnGluThrLeu 162
DB 1287 CGTATCGGACGCGCGCGCGATGACCGGAAATTTGATGACAGACGATCAGCAGTATC 1228
QY 163 ThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePheLeu 182
DB 1227 GCGGAA-----GTTGGTGTCTTTCGCGCTT 1201
QY 183 ArgGly-----IleAspValValSerProAlaAsnAlaAspThr 195
DB 1200 GAAGGAAAGCGGCTCATCGATCGATGATGATGCGTGAAGCCGCGCAATCGAGATA 1141
QY 196 AspValPheIleAsnAlaAspValPheGlyThrIleArgAsnArgThrGluMetHisLeu 215
DB 1140 CCCCTCTC-----GAGATCGAATCTC 1120
QY 216 TyrAsnAlaGluThrLeuValAlaGlnThrIleValLeuGluTyrPhe-----AlaValAsp 233
DB 1119 CCAGAGGCGGTCTCTGCGCATCAGCGCGCGCTGCAAGCGGTTGGCAAGCGCGCAGC 1060
QY 234 ArgThrAsnLysLysLeuLeuIleLysProLysThrAsnAlaPheGluAlaTyrLys 253
DB 1059 CGC-----GCCTTGAAGATGGGCAAGCTGCTTTCAGAGAGGCGCGCGGCTCAG 1006
QY 274 ---GlyLeuMet 276
DB 1005 CTCGCGCTCGTC 994

RESULT 12
US-08-426-630-29/c
Sequence 29, Application US/08426630
Patent No. 6656709
GENERAL INFORMATION:
APPLICANT: BLANCHE, FRANCIS; CAMERON, BEATRICE; CROUZET,
APPLICANT: JOEL; DEBUSCHE, LAURENT; LEVEY SCHIL, SOPHIE;
APPLICANT: THIBAUT, DENIS
TITLE OF INVENTION: POLYPEPTIDES INVOLVED IN THE
TITLE OF INVENTION: BIOSYNTHESIS OF COBALAMINS AND/OR COBAMIDES, DNA SEQUENCES
TITLE OF INVENTION: CODING FOR THESE POLYPEPTIDES, PREPARATION METHOD AND THEIR
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & PINNEGAN
STREET: 555 13TH STREET, N.W.
CITY: WASHINGTON
STATE: DISTRICT OF COLUMBIA
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426.630
FILING DATE:
CLASSIFICATION: 435

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,151
FILING DATE: 14-SEP-1992
APPLICATION NUMBER: PCT/FR91/00054
FILING DATE: 30-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: F. F. CALVERTI
REGISTRATION NUMBER: 28,557
REFERENCE/DOCKET NUMBER: 1290-7213
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 857-7887
TELEFAX: (202) 857-7929
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 4748 base pairs
TYPE: Nucleic Acid
STRANDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Pseudomonas denitrificans
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE:
ORGANELLE:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: Nucleotide sequence of the 5' to 3'
Patent No. 6656709
OTHER INFORMATION: strand from the 4748 bp SalI-SalI-SalI-SalI-BglI fragment
US-08-426-630-29
US-08-426-630-4 (1-320) x US-08-426-630-29 (1-4748)

Alignment Scores:
Pred. No.: 3.65 Length: 4748
Score: 95.00 Matches: 73
Percent Similarity: 31.10% Conserves: 34
Best Local Similarity: 21.22% Mismatches: 99
Query Match: 5.78% Indels: 138
DB: 4 Gaps: 17

QY 14 PheIleLeuSerAlaCyeGlyThrLeuThrGlyIleProSerHisGlyGlyLysArg 33
DB 4469 TTCGCTCGGAAGATGCGCGCGAGCTGTCGCT-----CATCGCGCGGCAAGCTC 4419
QY 34 PheAlaValGluGln---GluLeuValAlaAlaSerAlaArgAlaAlaValLysAspMet 52
DB 4418 ATCGCATCGACGATGTGACGCGACGCGCATAGTAGCGGTACAGTGTGCGCATACC 4359
QY 53 AspleuGlnAlaLeuHisGlyArg-----LysVal 62
DB 4358 GATCGCGCAGAGTTCACCGCGGAGCGGCTTTCGATCTGTCGATGACCGCGCGAGGTG 4299
QY 63 AlaLeuTyrIleAlaThrMetCglyAspGlnLyseryLysLeuThrGlyGlyArgTyr 82
DB 4298 ACGCTCCAGATGATCTCTGATTCACGA-----CAGGTGCA 4260
QY 83 SerIleAsp---AlaLeuIleArgGlyGluTyrIleAsnSerProAla----- 97
DB 4259 GTCTGACCGCGCGCGCGCTCCGAATCATCATCAGATGCGCGCTTTCGCGCGCGC 4200
QY 98 -----ValArgThrAspTyrThr 103
DB 4199 CATGACCGCTCATGCGCCCAATCAACGCTCGCGCTGATGTTTCTTTCGACGAGCC 4140

```

```
Qy 104 TyrProArgTyrGluThr----- 109
Db 4139 TTCCTGATCATTCAGCGCGAATTGCTGTCGCCGCGCCACGAGCGTGGCAGACTT 4080
Qy 110 -----ThrAlaGluThrThrseryl----- 116
Db 4079 GTAGACGATGTGTGCAGAGTGGTGTGAGGACCGCGGGCGTGTGGCTTCCGCGGACG 4020
Qy 117 -----GlyLeuThrGlyLeuThrThrsLeuSer 126
Db 4019 CCACTTCGCGTACGTCGCCACCTTCACGCTGTGGTAAACCGAGGATCTCGAC 3960
Qy 127 ThrLeuSnaIProAlaLeuSerArgThrGlnSerAsp-----GlySerGlySer 143
Db 3959 ---CTTACGCGCGACGCGCTGCAGCGTCCGCGCCAGATATCGCGCGAGGGCGGCGAC 3903
Qy 144 ArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAsp---TyrArgSnaIuThrLeu 162
Db 3902 CGTGATCGGACGCGCGCGATCGAGCGGAATTGTCATCGACAGCGTCACGACGGTATC 3843
Qy 163 ThrThrAsnProAlaGAspThrAlaPheLeuSerHisLeuValGlnThrValPhePheLeu 182
Db 3842 GCGGAA-----GTTGGTTCCTTTTCGCGCTT 3816
Qy 183 ArgGly-----IleAspValValSerProAlaAsnAlaAspThr 195
Db 3815 GAAGGAAGCGCGCTGCATCGGATCGATGATGATCGCTGAAGCCGCGCGGAATCGAGATA 3756
Qy 196 AspValPheIleAsnIleAspValPheGlyThrIleArgSnaIuThrGluMetHisLeu 215
Db 3755 CCCCTCTTC-----GAGATCGAACTC 3735
Qy 216 TyrTrsnaIaGluThrLeuValaGlnThrIleuGluTyrPhe-----AlaValAsp 233
Db 3734 CCAGAGCGGCTTCTGCTGCATCGAGCGCGCTGCAGCGGTTGGCAAGCGCGCCAC 3675
Qy 234 ArgThrAsnIleuValLeuLeuIleAspProIleThrAsnAlaPheGluAlaIleTyrIle 253
Db 3674 GCG----- 3672
Qy 254 GluAsnTyrAlaLeuTyrMetGlyProTyrIleValSerIleGlyIleLeuProThrGlu 273
Db 3671 -----GCCTTGAGATGGGCAAGCTGCTTCGAGGAAGGCGCGACGCGTGCAG 3621
Qy 274 ---GlyLeuMet 276
Db 3620 CTCGACCTCGTC 3609

RESULT 13
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
```

```
US-09-103-840A-2
Alignment Scores:
Pred. No.: 2.74e+05 Length: 4403765
Score: 95.00 Matches: 50
Percent Similarity: 36.68% Conservative: 23
Best Local Similarity: 25.13% Mismatches: 76
Query Match: 5.78% Indels: 50
DB: 3 Gaps: 10

US-10-617-835-4 (1-320) x US-09-103-840A-2 (1-4403765)
Qy 105 ProArgTyrGluThrThrAlaGluThrThrserylGlyLeuThrGlyLeuThrThrsSer 124
Db 315896 CCTAGCTGGAGGAAGCATGACATGACAA----- 315867
Qy 125 LeuSerThrLeuSnaIProAlaLeuSerArgThrGlnSerAspGlySerGlySerArg 144
Db 315866 CTGGAAATCCTCGCGACGCGACCGCGCTCGTCGAAAGACTCGGCGCGGACTG 315807
Qy 145 SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgSnaIuThr----- 161
Db 315806 GCCATCTCGGTGTGTCGTCATCGGTCGCGCGACCGCGCTCCACACCGTGGCCAC 315747
Qy 162 ---LeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db 315746 CCGCTAGTCGCGCAACCCGACGATTTGGCC-----ACGTCGAAGTGAACCTTC 315699
Qy 181 -----PheLeuArgGlyIleAspVal---ValSerProAlaAsnAlaAsp 194
Db 315698 GCGCGCTTCGCGCGCGGCTTGGCGGCGGACGTCACATCGCGGTGAGCGCGCGAC 315639
Qy 195 ThrAspValPheIleAsnIleAspValPheGlyThrIleArgSnaIuThrGluMetHis 214
Db 315638 ACCGATCCGACCGCGAAGCGATGATGTTGGACCC-----AACAGATTACACAC 315588
Qy 215 LeuTyrAsnAlaGluThrLeuValaGlnThrIle-----LeuGluTyrPhe 230
Db 315587 GTCCGGACGCGCGCATGATCTCACTGGGACCCACGCGCGCGGCTCGGACCTATCTG 315528
Qy 231 AlaValAspArgThrAsnIleuValLeuLeuIleAspProIleThrAsnAlaPheGluAla 250
Db 315527 GCGGTG-----CGCGCGCGGCTGTGTGGAGCC----- 315498
Qy 251 AlaTyrIleGluAsnTyrAlaLeuTyrMetGlyProTyrIleValSerIleGlyIleLeu 270
Db 315497 -----GTGCTGGGTCAAGGACTACGACGTCGATGTCGCGCATCGGC 315456
Qy 271 ProThrGluGlyLeuMetValAspPheSerAspIleArgProTyrGlyAsnHisThr 289
Db 315455 CCGTCGCG-----CTACGCGCGGAGACGTCGTCGCGCGGCGAAGACAC 315408

RESULT 14
US-09-071-035-157
; Sequence 157, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: GIL H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
```



```
Qy 154 MetGlyAspTyrArgAsnGluThrLeuThrThrAsnProArgAspThrAlaPheLeuSer 173
Db 700 GCCGGCTGATCACCAGACGAGCTGAGAACCGCCCT----- 738
Qy 174 HisLeuValGlnThrValPhePheLeuArgGlyIleAspValValSerProAlaAsnAla 193
Db 739 -----GAGTGGTCTCGGGCGGGCGGCC 762
Qy 194 AspThrAspValPheIleAsnIleAspValPheGlyThrIleArgAsnArgThrGluMet 213
Db 763 -----GCGACCTTCGCGAAACCGCCAGGCT 789
Qy 214 HisLeuTyrAsnAlaGluThrLeuLysAlaGlnThrLysLeu----- 227
Db 790 GCGCGCTATGCCGGGCAAGACCTCCGCGCTCAGGCCGAAAGCCCGGCGCTACCGGATCGTC 849
Qy 228 -----GluTyrPheAlaValAspArgThrAsnLysLysLeuIleLysPro 243
Db 850 GAGAACTCGACGAGCTGAAGCGCTGCGCGCCGCAACCAAG-----CAGCCG 900
Qy 244 LysThrAsnAlaPheGluAlaIleTyrLysGluAsnTyrAlaLeuTyrMetGlyProTyr 263
Db 901 CTGATCGGCTGTTC-----GCCCGGGCAACATGCCAGTGCCTGGCTCGGTCCGACC 954
Qy 264 LysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSerAspIleArg 283
Db 955 GCCACCTACCAAGC-----AACCTGAA- 977
Qy 284 ProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerHisGlu 303
Db 978 CCAGCCGGC-----GGTGAAGCTGCGAGCGCAACCGCAAGCGCACCGCGACATCCC 1028
Qy 304 GlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGln 319
Db 1029 GACCTGGCGCAATGACCAAGCCATCGAGCTGCTGAAGGACAA 1076
```

Search completed: August 19, 2005, 01:37:11  
Job time : 1983 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 18, 2005, 23:49:20 ; Search time 149 Seconds  
(without alignments)  
3514.151 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 320  
Sequence: 1 MRARLLRILFSLVFLSACG.....SHEGEGYDEAVRQHRCQP 320

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2400006

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cg2\_1/USPTO.spool.p/US10617835/runat\_18082005\_115731\_8766/app.query.fasta\_1.519  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2ncl1go.rn1 -MINMATCH=0.1  
-LOOPTCL=0 -LOOPTXT=0 -UNITS=bite -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdt -LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10617835\_@CGN\_1\_1\_105\_@runat\_18082005\_115731\_8766  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents\_NA:\*  
1: /cg2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cg2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cg2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cg2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cg2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cg2\_6/ptodata/1/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	320	100.0	1136	4	US-09-043-302-6
2	320	100.0	3287	4	US-09-043-302-1
3	8	2.5	38	4	US-09-371-772B-8146
4	8	2.5	38	4	US-09-371-772B-8900
5	8	2.5	38	4	US-09-371-772B-13757
6	8	2.5	1398	4	US-09-513-999C-16838
7	8	2.5	499	3	US-09-004-838-9
8	8	2.5	601	4	US-09-949-016-86885
9	8	2.5	601	4	US-09-949-016-105293
10	8	2.5	601	4	US-09-949-016-105294
11	8	2.5	601	4	US-09-949-016-126952
12	8	2.5	601	4	US-09-949-016-159911

C 13	8	2.5	601	4	US-09-949-016-159932	Sequence 159932.
C 14	8	2.5	636	4	US-09-902-540-8255	Sequence 8255, Ap
C 15	8	2.5	897	4	US-09-107-532A-72	Sequence 72, Appl
C 16	8	2.5	958	4	US-09-270-767-14224	Sequence 14224, A
C 17	8	2.5	966	4	US-09-252-991A-757	Sequence 757, App
C 18	8	2.5	1251	4	US-09-489-039A-3927	Sequence 3927, Ap
C 19	8	2.5	1848	1	US-08-333-802-1	Sequence 1, Appl1
C 20	8	2.5	2058	4	US-09-543-681A-310	Sequence 310, App
C 21	8	2.5	2499	4	US-09-252-991A-696	Sequence 696, App
C 22	8	2.5	2788	3	US-09-347-87B-45	Sequence 45, Appl
C 23	8	2.5	2871	4	US-09-614-221A-506	Sequence 506, App
C 24	8	2.5	3026	3	US-08-981-729-6	Sequence 6, Appl1
C 25	8	2.5	3026	4	US-09-613-811-6	Sequence 719, App
C 26	8	2.5	4519	4	US-09-949-016-719	Sequence 5111, App
C 27	8	2.5	4519	4	US-09-949-016-5111	Sequence 5111, Ap
C 28	8	2.5	4637	3	US-09-921-017B-818	Sequence 818, App
C 29	8	2.5	6492	4	US-09-902-540-853	Sequence 853, App
C 30	8	2.5	7282	4	US-09-949-016-17144	Sequence 17144, A
C 31	8	2.5	16056	4	US-09-949-016-16211	Sequence 16211, A
C 32	8	2.5	16056	4	US-09-949-016-16212	Sequence 16212, A
C 33	8	2.5	26086	4	US-09-949-016-15355	Sequence 15355, A
C 34	8	2.5	26238	4	US-09-949-016-15356	Sequence 12314, A
C 35	8	2.5	40352	3	US-08-846-111D-15	Sequence 15, Appl
C 36	8	2.5	40352	3	US-09-443-077-15	Sequence 14356, A
C 37	8	2.5	41617	4	US-09-949-016-14356	Sequence 14681, A
C 38	8	2.5	41618	4	US-09-949-016-14681	Sequence 7, Appl1
C 39	8	2.5	44377	2	US-08-804-222C-7	Sequence 1, Appl1
C 40	8	2.5	44377	2	US-08-804-198-1	Sequence 12461, A
C 41	8	2.5	93920	4	US-09-949-016-12461	Sequence 16853, A
C 42	8	2.5	93920	4	US-09-949-016-16853	Sequence 16853, A
C 43	8	2.5	247781	4	US-09-949-016-14193	Sequence 14193, A
C 44	8	2.5	247781	4	US-09-949-016-14193	Sequence 14193, A
C 45	8	2.5	536165	4	US-09-214-808-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-043-302-6  
Sequence 6, Application US/09043302  
Patent No. 6617128  
GENERAL INFORMATION:  
APPLICANT: MEYER, Thomas F.  
APPLICANT: RUDEL, Thomas  
APPLICANT: SCHEUERPLUG, Ina  
APPLICANT: WALTER, Jurgen  
APPLICANT: EICKERJAGER, Sandra  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
TITLE OF INVENTION: the Adhesion of Neisseria Cells to Human Cells  
FILE REFERENCE: 0147-172P  
CURRENT APPLICATION NUMBER: US/09/043,302  
CURRENT FILING DATE: 1998-06-08  
EARLIER APPLICATION NUMBER: PCT/EP96/04092  
EARLIER FILING DATE: 1995-09-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 1136  
TYPE: DNA  
ORGANISM: Neisseria gonorrhoeae  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (135)..(1094)  
US-09-043-302-6  
Alignment Scores:  
Pred. No.: 0  
Score: 320.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Length: 1136  
Matches: 320  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-617-835-4 (1-320) X US-09-043-302-6 (1-1136)

Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
Db 135 ATGGGGGACGGCTGATACCTATTCTTTTTCGGTTTATTTATTCGGCCGCGG 194  
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGluLeu 40  
Db 195 ACACGTACAGATATTCATCCGATGCGGAGGCAACCTTCGCGGTGAAACAAGACTT 254  
Qy 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
Db 255 GTGGCCGCTTCGCGACAGCTGCGGTAAAGCATGATTTACAGGCATTACAGGACGA 314  
Qy 61 LysValAlaLeuValIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGly 80  
Db 315 AAAGTTGATTTGATTCATTGCAACTATGGGCGACCAAGGTTGACGAGTTCACAGGGGGT 374  
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100  
Db 375 CGCTACTCCATTGATGCACTGATTCGGCGCAATACAAACAGCCCTCGCTCGCAC 434  
Qy 101 AspTyrThrTyrProArgTyrGluThrThraIaGluThrThrSerGlyGlyLeuThrGly 120  
Db 435 GATTACACCTATCCGCGTTACGAAACCAACCGCTGAAACATCAGCGCGTTGACGGGT 494  
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
Db 495 TTAAACCACTCTTATATCTACCTTAATGCCCCCTGACCTTGCCGACCAACCAATCGACGGT 554  
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
Db 555 AGCGGAAGTAGAGACAGCTCGGCTTAATTTGGCGGGAATGGGGGANTTACGAATGAA 614  
Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180  
Db 615 ACCTTGACGACCAACCGCGCGACACTGCTTCTTCCCTTCCCTGTTACACACCGTATTT 674  
Qy 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
Db 675 TTCCGCGCGGACATGACGATGCTTCTCTCCCAATGCGCATACGATGCTTATTAAC 734  
Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220  
Db 735 ATGACGATATTCGAGATGATACGACACGAAACCGAATGACCTATACAAATGCCGAACA 794  
Qy 221 LeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLysLeuLeu 240  
Db 795 CTGAAACCCCAAAACAACTGGAATATTTCCAGTAGACAGAACCAATAAAAAATTGCTC 854  
Qy 241 IleLysProLysThrAsnAlaPheGluAlaIleAlaTyrLysGluAsnTyrAlaLeuTyrMet 260  
Db 855 ATCAAAACCCCAAAACCAATGCTTTGAAGCTGCTTAAAGAAATTTACGACTTTGGATG 914  
Qy 261 GlyProTyrLysValSerLysGlyIleLysProThrGluGlyLeuMetValAspPheSer 280  
Db 915 GGGCGGATATAAGTAAGCAAAAGAAATCAACCGACGGAAGGATTAATGTCGATTTCTCC 974  
Qy 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsn 300  
Db 975 GATATTCGGGCATACGGCAATCATACGGGTAACTCCGCCCATCCGAGAGGCGAATAAC 1034  
Qy 301 SerHisGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyPro 320  
Db 1035 AGTCATAGGGGTATGATACAGCGATGACGAGTTCGACACATTAACAAAGGCAACT 1094

RESULT 2  
US-09-043-302-1  
; Sequence 1, Application US/09043302; Patent No. 6617128  
; GENERAL INFORMATION:  
; APPLICANT: MEYER, Thomas F.

APPLICANT: RUDEL, Thomas  
APPLICANT: SCHERERPLUG, Ina  
APPLICANT: MATER, Jurgen  
APPLICANT: EICKERUDAGER, Sandra  
APPLICANT: SCHWAN, Thomas  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
FILE REFERENCE: 0147-172P  
CURRENT APPLICATION NUMBER: US/09/043,302  
EARLIER FILING DATE: 1998-06-08  
EARLIER APPLICATION NUMBER: PCT/EP96/04092  
EARLIER FILING DATE: 1995-09-18  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3287  
TYPE: DNA  
ORGANISM: Neisseria gonorrhoeae  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (136)..(447)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (583)..(1542)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1585)..(3111)  
US-09-043-302-1

Alignment Scores:  
Pred. No.: 0 Length: 3287  
Score: 320.00 Matches: 320  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-617-835-4 (1-320) X US-09-043-302-1 (1-3287)

Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
Db 583 ATGGGGGACGGCTGATACCTATTCTTTTTCGGTTTATTTATTCGGCTCGCGG 642  
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyLysArgPheAlaValGluGluLeu 40  
Db 643 ACACGTACAGATATTCATGCACTGCGGAGGCAAAAGCTTCGGCGTGAACAAGACTT 702  
Qy 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
Db 703 GTGGCCGCTTCGCGACAGCTGCGCTTAAAGACATGATTTACAGCATTAACAGGACGA 762  
Qy 61 LysValAlaLeuValIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGly 80  
Db 763 AAAGTTGATTTGATTCATTCGACTATGCGGACCAACAGTTCAGGCACTTTGACAGGGGT 822  
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100  
Db 823 CGCTACTCCATTGATGCACTGATTCGGCGCAATACATAAAGCCCTGCGCTCGCAC 882  
Qy 101 AspTyrThrTyrProArgTyrGluThrThraIaGluThrThrSerGlyGlyLeuThrGly 120  
Db 883 GATTACACCTATCCGCTTACGAAACCAACCGCTGAAACATCAGCGGTTGACGGGT 942  
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
Db 943 TTAAACCACTCTTATATTAACCTTAATGCCCTGCACTTCGCGCAACCATCAGAGGT 1002  
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
Db 1003 AGCGGAAGTAGAGACAGCTCGGCTTAATTTGCGCGGATGGGGGATTTATCGAAATGAA 1062  
Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180

Db 1063 ACCTGACGACACCGCGGCGACACTGCTTTCTTCCACTGTGTAACAGCCGATTT 1122  
Qy PheLeuAaGgLYlLaSpVaVaISeRPoAlaAspThRaSpVaIheIleAen 200  
Db 1123 TTCCGCGCGGCGATAGACGTTGTTCTCTCCCAATGCGATACGATGTTTATTAC 1182  
Qy 201 ILeaSpVaIaPheGlyThIleAaSpAaRgThRgImeHleuTyRaSpAlaGluPhr 220  
Db 1183 ATCGACGATTTGCAACGATACGACACGACCAATGACCTTATACATGCCAACA 1242  
Qy 221 LeuLyAaGInThRyLeuGluTyRheaIaVaIaSpAaRgThRaSpLyLeuLeu 240  
Db 1243 CTGAAGCCCAAAACAACTGGAATATTTCGACAGACAGAACCAATAAAAATTGCTC 1302  
Qy 241 ILeLySPoLyThRaSpAlaPheGluAlaAlaTyLySGluAsnTyRaLeuTPhet 260  
Db 1303 ATCAAAACCCAAACCAATGCGTTTGAAGCTGCTTAAAGAAATTTACGCAATTGGATG 1362  
Qy 261 GLYProTyRyLyVaISeRySGlyIleLySPoThRgIleuMeTyVaIaSpPheSer 280  
Db 1363 GGGCGGTATTAAGTAAGCAAGCAATCAACCGACGAAAGATTAAATGTCGATTTCTCC 1422  
Qy 281 AspIeaRgProTyRyGlyAaSnIeThRyAaSeRaIaProSeVaIgluAlaAaSpAa 300  
Db 1423 GATATCCGCGCATACGCGCATATACGGGTACTCGCCCATCCGTAGAGCGTATAC 1482  
Qy 301 SerHISgluGlyTyRyGlyTySeRaSpGluAlaVaIaRgGlnHISaRgIngluInPro 320  
Db 1483 AGTCATGAGGGGTATGATACGATGCAAGCAAGTGCACAACTATGACGAGGCAACT 1542

## RESULT 3

US-09-371-772B-8146  
; Sequence 8146, Application US/09371772B  
; Patent No. 6566127  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: McSwiggen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00, 876-J (237/198)  
; CURRENT APPLICATION NUMBER: US/09/371,772B  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 14225  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8146  
; LENGTH: 38  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid  
US-09-371-772B-8146

Alignment Scores:  
Pred. No.: 4.57 Length: 38  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.50% Indels: 0  
DB: 4 Gaps: 0

US-10-617-835-4 (1-320) x US-09-371-772B-8146 (1-38)

Qy 306 GIIYTySeRaSpGluAlaVaIaRg 313  
Db 2 GGUUACUCUGAUGAGCGCGUUGAG 25

RESULT 4  
US-09-371-772B-8900  
; Sequence 8900, Application US/09371772B  
; Patent No. 6566127  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: McSwiggen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00, 876-J (237/198)  
; CURRENT APPLICATION NUMBER: US/09/371,772B  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 14225  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 8900  
; LENGTH: 38  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid  
US-09-371-772B-8900

Alignment Scores:  
Pred. No.: 4.57 Length: 38  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.50% Indels: 0  
DB: 4 Gaps: 0

US-10-617-835-4 (1-320) x US-09-371-772B-8900 (1-38)

Qy 306 GIIYTySeRaSpGluAlaVaIaRg 313  
Db 2 GGUUACUCUGAUGAGCGCGUUGAG 25

RESULT 5  
US-09-371-772B-13757  
; Sequence 13757, Application US/09371772B  
; Patent No. 6566127  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: McSwiggen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00, 876-J (237/198)  
; CURRENT APPLICATION NUMBER: US/09/371,772B  
; CURRENT FILING DATE: 1999-08-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 14225  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13757  
; LENGTH: 38  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Enzymatic Nucleic Acid  
NAME/KEY: misc\_feature  
LOCATION: (31)..(31)  
; OTHER INFORMATION: n stands for inosine

```

US-09-371-772B-13757
Alignment Scores:
Pred. No.: 4.57 Length: 38
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 4 Gaps: 0

US-10-617-835-4 (1-320) x US-09-371-772B-13757 (1-38)
Qy 306 GlyTyrSerAspGluAlaValArg 313
Db 2 GGCACUCUGAGUGAGCCGCUAGG 25

RESULT 6
US-09-513-999C-16838
; Sequence 16838, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 16838
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 21
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 40
; OTHER INFORMATION: d=a or g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 326
; OTHER INFORMATION: n=a, g, c or t
; US-09-513-999C-16838

Alignment Scores:
Pred. No.: 51.9 Length: 398
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 4 Gaps: 0

US-10-617-835-4 (1-320) x US-09-513-999C-16838 (1-398)
Qy 116 GlyGlyLeuThrGlyLeuThrThr 123
Db 140 GGAGATTGACAGGTTTGACAACT 163

RESULT 7
US-09-004-838-9
; Sequence 9, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelson, Richard W.
; APPLICANT: Shen, Kathy
; APPLICANT: Meyers, Blake

TITLE OF INVENTION: Procedures and Materials for
TITLE OF INVENTION: Confering Pest Resistance in Plants
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,838
FILING DATE: 09-JAN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/781,734
FILING DATE: 10-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 023070-078810US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: -
LOCATION: 1..499
OTHER INFORMATION: /note="RGLIT"

US-09-004-838-9

Alignment Scores:
Pred. No.: 65.5 Length: 499
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.50% Indels: 0
DB: 3 Gaps: 0

US-10-617-835-4 (1-320) x US-09-004-838-9 (1-499)
Qy 12 SerValPheIleuSerAlaCys 19
Db 379 TCTGCTTATTCTGTCAGCATGC 402

RESULT 8
US-09-949-016-86885/C
; Sequence 86885, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 86885
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-86885

Alignment Scores:
Pred. No.: 79.4      Length: 601
Score: 8.00          Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4                Gaps: 0

US-10-617-835-4 (1-320) x US-09-949-016-86885 (1-601)

Qy 74 SerGlySerLeuThrGlyGlyArg 81
Db 570 TCCGGCTCAGTTCAGGTCGAGG 547

RESULT 9
US-09-949-016-105293/c
; Sequence 105293, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 105293
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105293

Alignment Scores:
Pred. No.: 79.4      Length: 601
Score: 8.00          Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4                Gaps: 0

US-10-617-835-4 (1-320) x US-09-949-016-105293 (1-601)

Qy 15 IleuSerAlaCysGlyThrLeu 22
Db 418 ATTCTCAGTCGCTGTCGATCATTA 395

RESULT 10
US-09-949-016-105294/c
; Sequence 105294, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 105294
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105294

Alignment Scores:
Pred. No.: 79.4      Length: 601
Score: 8.00          Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4                Gaps: 0

US-10-617-835-4 (1-320) x US-09-949-016-105294 (1-601)

Qy 15 IleuSerAlaCysGlyThrLeu 22
Db 68 ATTCTCAGTCGCTGTCGATCATTA 45

RESULT 11
US-09-949-016-126952
; Sequence 126952, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 126952
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-126952

Alignment Scores:
Pred. No.: 79.4      Length: 601
Score: 8.00          Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4                Gaps: 0

US-10-617-835-4 (1-320) x US-09-949-016-126952 (1-601)

Qy 141 SerGlySerArgSerSerLeuGly 148
Db 373 AGTGGTCGAGATCCAGCTGGGT 396

RESULT 12
US-09-949-016-159911/c
; Sequence 159911, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 159911
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-159911

Alignment Scores:
Pred. No.: 79.4      Length: 601
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4      Gaps: 0

US-10-617-835-4 (1-320) x US-09-949-016-159911 (1-601)

QY 140 GlySerGlySerArgSerSerLeu 147
Db 197 GGGAGCGGAGCAGGTCTCCCTC 174

RESULT 13
US-09-949-016-159932/c
; Sequence 159932, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 159932
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-159932

Alignment Scores:
Pred. No.: 79.4      Length: 601
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4      Gaps: 0

US-10-617-835-4 (1-320) x US-09-949-016-159932 (1-601)

QY 140 GlySerGlySerArgSerSerLeu 147
Db 197 GGGAGCGGAGCAGGTCTCCCTC 174

RESULT 14
US-09-902-540-8255
```

```

; Sequence 8255, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8255
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8255

Alignment Scores:
Pred. No.: 84.2      Length: 636
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 2.50%      Indels: 0
DB: 4      Gaps: 0

US-10-617-835-4 (1-320) x US-09-902-540-8255 (1-636)

QY 41 ValAlaAlaSerAlaArgAlaAla 48
Db 270 GTGCGGCGTTCGCTCGAGCTGCC 293

RESULT 15
US-09-107-532A-72
; Sequence 72, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinelli, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 897 base pairs
```

```

:      TYPE: nucleic acid
:      STRANDEDNESS: double
:      TOPOLOGY: circular
:      MOLECULE TYPE: DNA (genomic)
:      HYPOTHETICAL: NO
:      ANTI-SENSE: NO
:      ORIGINAL SOURCE:
:      ORGANISM: Enterococcus faecium
:      FEATURE:
:      NAME/KEY: misc_feature
:      LOCATION: (B) LOCATION 1...897
:      SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-107-532A-72

```

```

Alignment Scores:
Pred. No.:      120      Length:      897
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.50%      Indels:      0
DB:                4      Gaps:          0

```

US-10-617-835-4 (1-320) x US-09-107-532A-72 (1-897)

```

Oy      14  PheileusevAlaCyagLYThr 21
          |||||||
Db      103 TTATTTTAAGCGCTTGCGGTACA 126

```

Search completed: August 19, 2005, 01:21:23  
 Job time : 160 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 19, 2005, 01:39:13 ; Search time 650 Seconds

(without alignments)  
3198.463 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 320

Sequence: 1 MRARLLPILFVFLSACG.....SHEGYSDAVRQRQGP 320

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 7316285 seqs, 324859403 residues

Word size: 1

Total number of hits satisfying chosen parameters: 14620882

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp

-O=/cgn2\_1/USPTO.epool.p/US10617835/runat.18082005.115732.8783/app.query.fasta.1.519

-DB=Published Applications NA -QMT=fastcap -SUFFIX=p2noligo.rnpb -MINMATCH=0.1

-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=oligo

-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1

-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USER=US10617835.OCGN.1.1.723.0/runat.18082005.115732.8783

-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*

6: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*

8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*

9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*

10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*

11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*

13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

17: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

18: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

19: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

20: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

21: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

22: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

23: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

24: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

25: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

26: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	320	100.0	1136	US-09-043-302-6	Sequence 6, Appl1
2	320	100.0	1136	US-10-617-835-6	Sequence 6, Appl1
3	320	100.0	1136	US-09-043-302-1	Sequence 1, Appl1
4	320	100.0	1136	US-10-617-835-1	Sequence 1, Appl1
5	320	100.0	1136	US-10-617-835-1	Sequence 6, Appl1
6	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
7	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
8	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
9	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
10	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
11	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
12	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
13	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
14	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
15	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
16	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
17	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
18	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
19	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
20	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
21	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
22	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
23	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
24	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
25	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
26	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
27	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
28	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
29	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
30	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
31	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
32	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
33	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
34	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
35	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
36	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
37	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
38	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
39	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
40	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
41	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
42	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
43	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
44	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1
45	320	100.0	1136	US-10-617-835-1	Sequence 253306, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-043-302-6  
; Sequence 6, Application US/09043302  
; Publication No. US20020086349A1  
GENERAL INFORMATION:  
APPLICANT: MEYER, Thomas F.  
APPLICANT: RUDEL, Thomas  
APPLICANT: SCHERER, Ina  
APPLICANT: MAIER, Jürgen  
APPLICANT: EICKENJÄGER, Sandra  
APPLICANT: SCHMANN, Thomas  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart  
FILE REFERENCE: 0147-172P  
CURRENT APPLICATION NUMBER: US/09/043.302  
CURRENT FILING DATE: 1998-06-08

Pred. No. is the number of results predicted by chance to have a

```

; EARLIER APPLICATION NUMBER: PCT/EP96/04092
; EARLIER FILING DATE: 1995-09-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 6
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1094)
; US-09-043-302-6

Alignment Scores:
Pred. No.: 0 Length: 1136
Score: 320.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-617-835-4 (1-320) x US-09-043-302-6 (1-1136)
QY 1 MetAaGAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db 135 ATGGGGGACAGGCTGCTGATACCTATCTTTTTCGGTTTATTTATCCGCCCTGGCGG 194
QY 21 ThrLeuThrglyIleProSerHisGlyGlyGlyValAspPheAlaValGluGluLeu 40
Db 195 ACACGACAGGATTCATTCGATGCGGAGGCAAAACCTTCGCGGTGAACAGAACTT 254
QY 41 ValAlaAlaSerAlaArgAlaAlaValIysAspMetAspLeuGluAlaLeuHisGlyArg 60
Db 255 GTGGCCGCTTTCGCGAGAGCTGCCGTTAAAGCATGATTTACAGGCATTTACAGGACGA 314
QY 255 GTGGCCGCTTTCGCGAGAGCTGCCGTTAAAGCATGATTTACAGGCATTTACAGGACGA 314
Db 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrglyGly 80
Db 315 AAAGTTGATGTGATTCATTGCAACTATGGCGACCAAGGTTCAAGCATTTGACAGGGGGT 374
QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrglyGly 80
Db 81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100
Db 375 CGCTACTCATTTGATGACTGATTCGCGGCAATACATTAACAGCCCTGCGCTCGCACCC 434
QY 101 AspTyrThrglyProArgTyrGluThrThraIleGluThrThraSerGlyLeuThrgly 120
Db 435 GATTACACCTATCCGCGTTACGAACCAACCGCTGAACACATCAGCGGTTTGAAGGGGT 494
QY 121 LeuThrThrglySerLeuSerThraLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db 495 TTAAACCACTTCCTTAATCTACACTTAATCCCTGCACTCTGCGCACCAATCAAGACGGT 554
QY 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160
Db 555 AGCGGAAGTACGACGATCTGGGCTTAATTTTGGCGGGAATGGGGATTAACGAATGAA 614
QY 161 ThrLeuThrThraAsnProArgAspThraIlePheLeuSerHisGlyValGlnThrValPhe 180
Db 615 ACCTTGAGACCAACCCGCGGACACTGCTTTTCCCACTGGTACAGACGTAATTT 674
QY 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db 675 TTCCTGCGGCAATAGAGGTTGTTCTCTCCGCAATGCGCATACAGATGTTTAATTAAC 734
QY 201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisGlyTyrAsnAlaGluThr 220
Db 735 ATCGACGATTCGAGACGATACGCAACGAACCGAAATGCACTTACATACAGCCGAACA 794
QY 221 LeuIysAlaGlnThrIysLeuGluIuTyrPheAlaValAspArgThrAsnIleGlyLeu 240
Db 795 CTGAAGAGCCCAAACTGGAATATTTCCAGATGACAGAACCAATTAATAAATTTGGTCC 854
QY 241 IleLysProLysThrAsnAlaPheGluAlaAlaIleTyrIleAsnSerProAlaValArgThr 260

```

```

Db 855 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTTAAAGAAATTAACGATTTGTGATG 914
QY 261 GlyProTyrIysValSerIysGlyIleIysProThrglyLeuMetValAspPheSer 280
Db 915 GGGCGGATTAAGTAAAGTACGAAGGATCAACCGAGGAGGATTAAGTCGATTTCTCC 974
QY 281 AspIleArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluIleAspAsn 300
Db 975 GATATCCGGCCATACGCGCAATCATATGAGTAACTCCGCCCATTCGTAGAGGCTGATTAAC 1034
QY 301 SerHisGlyGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyPro 320
Db 1035 AGTCATGAGGGGTATGATACAGCATGAGACAGTGGCAACATGACAAAGGGCAACT 1094

RESULT 2
US-10-617-835-6
; Sequence 6, Application US/10617835
; Publication No. US20050124037A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHEUERPLUG, Ina
; APPLICANT: MATER, Jürgen
; APPLICANT: EICKENHÄGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; FILE REFERENCE: 0147-172P
; CURRENT APPLICATION NUMBER: US/10/617,835
; PRIOR FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US/09/043,302
; PRIOR FILING DATE: 1998-06-08
; PRIOR APPLICATION NUMBER: PCT/EP96/04092
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 6
; LENGTH: 1136
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1094)
; US-10-617-835-6

Alignment Scores:
Pred. No.: 0 Length: 1136
Score: 320.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-617-835-4 (1-320) x US-10-617-835-6 (1-1136)
QY 1 MetAaGAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db 135 ATGGGGGACAGGCTGCTGATACCTATCTTTTTCGGTTTATTTATCCGCCCTGGCGG 194
QY 21 ThrLeuThrglyIleProSerHisGlyGlyGlyValAspPheAlaValGluGluLeu 40
Db 195 ACACGACAGGATTCATTCGATGCGGAGGCAAAACCTTCGCGGTGAACAGAACTT 254
QY 41 ValAlaAlaSerAlaArgAlaAlaValIysAspMetAspLeuGluAlaLeuHisGlyArg 60
Db 255 GTGGCCGCTTTCGCGAGAGCTGCCGTTAAAGCATGATTTACAGGCATTTACAGGACGA 314
QY 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrglyGly 80
Db 315 AAAGTTGATGTGATTCATTGCAACTATGGCGACCAAGGTTCAAGCATTTGACAGGGGGT 374
QY 81 ArgTyrSerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThr 100

```

```

Db      375 CGCTACTCCATTGATGACTGATGCGGCGAATACATAACAGCCCTGCGCGACC 434
Qy      101 AaplyrthrrprxargtyrglunthrrhralagluThhSerGlyglyleuthrGly 120
Db      435 GATTACACCTTATCCGGTTACGAAACACCGCTGAAACAACTCAGGCGGTTGACGGGT 494
Qy      121 LeuthrrSerLeuSerThrLeuSnnalAProAlaLeuSerArgThrglnSerapgly 140
Db      495 TTAAACCACTTCTTTATCTACACTTAATGCCCTGCACCTTCCGCGACCAATCAAGCGT 554
Qy      141 SerGlySerArgSerSerleuGlyLeuSnnalleglyGlyMetGlyAapTyrrarganglu 160
Db      555 AGCGGAAGTAGAGAGCTGCTGGCTTAATATTTGGCGGATGGGGGATTAATGAAATGA 614
Qy      161 ThrLeuthrrThraenProargapThrrAlaPheleuSerHisleuValGlnThrValPhe 180
Db      615 ACCTTGACGACCAACCCGCGGACACTGCTTTCTTCCACTTGTCAGACCGTAATT 674
Qy      181 PheleuArgGlyIleAapValIalSerProAlaAaAalAapThrAapValPheIleAen 200
Db      675 TTCCTGCGGCGCATAGACGTTCTTCTCTGCCAATGCCGATACAGATGTTTATTTAC 734
Qy      201 IleapValIleaglyThrrIleargSnnargThrglumethIsleuTyrrAaAalagluThr 220
Db      735 ATCGACGATTCGGAACGATACGCAACGAAACCGAAATGCACCTATTAATGCCAACA 794
Qy      221 LeuysAlaaglnThrrlyleuGlnunrrPheAlaValIaapArgThrrAaAalyleuLeu 240
Db      795 CTGAAAGCCCAAAACAACTGGAATATTTCCGATGACAGAACCAATAAAAATTTGCTC 854
Qy      241 IleuysProLySthrrAaAalAphelGluAalAlyrlySgluAaAalaleuThrrPmet 260
Db      855 ATCAAAACCAAAACCAATGCCCTTGAAGCTGCTTAAAGAAATTAAGCATTTGCGATG 914
Qy      261 GlyProTyrrlySvalSerlySgllyIleuysProThrgluglyLeuMetValAapPheSer 280
Db      915 GGGCGCGTAAAGTAAGCAAGAAATCAAAACCGAAGCATTAATGATGATTTCTCC 974
Qy      281 AapIleAargProTyrrGlyAaAalAethrrGlyAaAalAaProSeValGluAalAapAen 300
Db      975 GATATCCGCGCATAGGCAATCATACGGGTAACTCCGCCCATCCGTAGAGGCTGATTAAC 1034
Qy      301 SerHisGlyGlyTyrrGlyTyrrSerAapGluAalAalAargGlnHisAargGlnPro 320
Db      1035 AGCTATGAGGGGTATGATACAGCAATGACGATGACCAATAGCAAGGCAACT 1094

RESULT 3
US-09-043-302-1
; Sequence 1, Application US/09043302
; Publication No. US20020086349A1
; GENERAL INFORMATION:
; APPLICANT: MEYER, Thomas F.
; APPLICANT: RUDEL, Thomas
; APPLICANT: SCHEUERPFUG, Ina
; APPLICANT: MAIER, Jurgen
; APPLICANT: EICKERNJAGER, Sandra
; APPLICANT: SCHWAN, Thomas
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Proteins Which Impart
; TITLE OF INVENTION: the Adhesion of Neisseria Cells to Human Cells
; FILE REFERENCE: 0147-172P
; CURRENT FILING DATE: 1998-06-08
; EARLIER FILING DATE: 1995-09-18
; EARLIER APPLICATION NUMBER: PCT/EP96/04092
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3287
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(447)

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (583)..(1542)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1585)..(3111)
US-09-043-302-1
Alignment Scores:
Pred. No.: 0
Score: 320.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9 Gaps: 0

US-10-617-835-4 (1-320) x US-09-043-302-1 (1-3287)
Qy      1 MetArgAlaArgLeuLeuIleProIleuPheSerValPheIleuSerAlaCySgly 20
Db      583 ATGCGGCGACGGCTGCTGATACCTATTTCTTTTCGTTTATTTATTCGCGCTGCGG 642
Qy      21 ThrLeuthrrGlyIleProSerHisGlyGlyLySargPheAlaValGlnGlnluLeu 40
Db      643 AACTGACAGGATTCATTCGATGCGGAGGCAAAACCTTCGCGGTGAAACAACTT 702
Qy      41 ValAlaAlaSerAlaArgAlaAlaValIyAapPheAapleuGlnAlaLeuHisGlyArg 60
Db      703 GTGGCGGCTTCTGCGACAGCTGCGGTTAAACATGATTAACAGCAATTACACGAGCA 762
Qy      61 LySValAlaLeuTyrrIleAlaThrrMetGlyAapGlnGlySerGlySerLeuThrrGly 80
Db      763 AAAGTGCATTTGATTCATTGCAACTATGCGGCAACCAAGTTCAAGCATTTGACAGGGGT 822
Qy      81 ArgTyrrSerIleAapAlaLeuIleargGlyGluTyrrIleAaAalSerProAlaValArgThr 100
Db      823 CGTACTCCATTGATGACTGATTCGCGGCAATACATTAACAGCCCTGCGCGCAC 882
Qy      101 AapTyrrThrrprxargtyrglunthrrhralagluThhSerGlyglyleuthrGly 120
Db      883 GATTACACCTTATCCGGTTACGAAACCAACCGCTGAAACAACTCAGGCGGTTTACGGGT 942
Qy      121 LeuthrrSerLeuSerThrLeuSnnalAProAlaLeuSerArgThrglnSerapgly 140
Db      943 TTAAACCACTTCTTTATCTACACTTAATGCCCTGCACCTTCCGCGACCAATCAAGCGT 1002
Qy      141 SerGlySerArgSerSerleuGlyLeuSnnalleglyGlyMetGlyAapTyrrarganglu 160
Db      1003 AGCGGAAGTAGAGAGCTGCTGGCTTAATATTTGGCGGATGGGGGATTAATGAAATGA 1062
Qy      161 ThrLeuthrrThraenProargapThrrAlaPheleuSerHisleuValGlnThrValPhe 180
Db      1063 ACCTTGACGACCAACCCGCGGACACTGCTTTCTTCCACTTGTCAGACCGTAATT 1122
Qy      181 PheleuArgGlyIleAapValIalSerProAlaAaAalAapThrAapValPheIleAen 200
Db      1123 TTCCTGCGGCGCATAGACGTTCTTCTCTGCCAATGCCGATACAGATGTTTATTTAAC 1182
Qy      201 IleapValIleaglyThrrIleargSnnargThrglumethIsleuTyrrAaAalagluThr 220
Db      1183 ATCGACGATTCGGAACGATACGCAACGAAACCGAAATGCACCTATTAACAAAGCCGAAACA 1242
Qy      221 LeuysAlaaglnThrrlyleuGlnunrrPheAlaValIaapArgThrrAaAalyleuLeu 240
Db      1243 CTGAAAGCCCAAAACAACTGGAATATTTCCAGATGACAGAACCAATAAAAATTTGCTC 1302
Qy      241 IleuysProLySthrrAaAalAphelGluAalAlyrlySgluAaAalaleuThrrPmet 260
Db      1303 ATCAAAACCAAAACCAATGCCCTTGAAGCTGCTTAAAGAAATTAAGCATTTGCGATG 1362
Qy      261 GlyProTyrrlySvalSerlySgllyIleuysProThrgluglyLeuMetValAapPheSer 280
Db      1363 GGGCGGTATAAAGTAAGCAAGAAATCAAAACCGAAGCATTAATGATGATTTCTCC 1422

```

Db	823	CGCTACTCCATTGATGACACTGATTCCGGCGGGAATACATTAAACAGCCCTGCGCTCGGACC	882
Oy	101	AspTyrThrTYrProAqTYrGLuThrThAlaGluThrThrsrGlyGlyLeuThrGly	120
Db	883	GATTACACCTATCCGCGCTTACGAAACACCGCGTGAACAACATCAGCGGTTTGAACGGGT	942
Oy	121	LeuThrThrsrLeuSerSerThrLeuAsnAlaProAlaLeuSerAqThrGlnSerAspGly	140
Db	943	TTAACCACTTCTTATCTACACTTAAATGCCCTCGCACTCTCGCGACCCCAATCAGACGCT	1002
Oy	141	SerGlySerAqrgsrSerSerLeuGlyLeuAsnAlaGlyGlyMetGlyAspTYrAqsrGlu	160
Db	1003	AGCGGAAGTGGAGACACTCTGGCGCTTAAATTTGGCGGATGGGGGATTTATCGAAATGAA	1062
Oy	161	ThrLeuThrThrsrProAqsrThrAlaPheLeuSerHisrLeuValGlnThrValPhe	180
Db	1063	ACCTTGACGACCAACCCGCGGAGACAGCTTCTTTCCTCCACTTGGTACAGCCGTAATTT	1122
Oy	181	PheLeuAqrgGlyLeaAspValAlaSerProAlaAsnAlaAspThraAspValPheLeaAsn	200
Db	1123	TTCTGTGGCGGCATAGACGTTGTTTCTTCCTGCCAATCCGATACAGATGTTTATTAAC	1182
Oy	201	IlleAspValPheGlyThrIleAqsrAsnAqThrGluMetHisrLeuTYrAsnAlaGluThr	220
Db	1183	ATCAGCACTATTCCGAAAGATACGCAACGAAACGAAATGCACTTATACATGCCGAACA	1242
Oy	221	LeuLYsAlaGlnThraTYrLYsLeuGluTYrPheAlaValAspAqrgThraenLYsLYsLeu	240
Db	1243	CTGAAGACCCCAACAAACCTGGAATATTTCGCGATACACGAACCAATAAAAATTCCTC	1302
Oy	241	IleLYsProLYsThraAsnAlaPheGluAlaAlaTYrLYsGluAsnTYrAlaLeuTYrMet	260
Db	1303	ATCAAAACCCAAACCAATGCGTTGAAAGCTGCGCTATGAACAAATTCGCAATGTGGATG	1362
Oy	261	GlyProTYrLYsValSerLYsGlyIleLYsProThrGluGlyLeuMetValaAspPheSer	280
Db	1363	GGGCGCGTAAAGTACGCAAAAGAAATCAAAACGACGAAAGATTAATGCTCGATTCTCC	1422
Oy	281	AspIleAqrgProTYrGlyAsnHisrThrGlyAsnSerAlaProSerValGluAlaAspAsn	300
Db	1423	GATATCCGGCATATCGGCATATCAATACGGGTAACTCCGCCCATCCGTAAGGCGCTGATTAAC	1482
Oy	301	SerHisrGluGlyTYrTYrTYrSerAspGluAlaValaArgGlnHisrAqrgGlnGlyAlaPro	320
Db	1483	AGTCATAGAGGGGTATGAGTACAGCATGAAGCACTGGCAACAATTAACAGGCGCAACT	1542
RESULT 5			
US-10-320-800-69			
Sequence 69, Application US/10320800			
Publication No. US20030215462A1			
GENERAL INFORMATION:			
APPLICANT: ROBINSON, ANDREW			
APPLICANT: GORRINGE, ANDREW			
APPLICANT: HUDSON, MICHAEL			
APPLICANT: REDDIN, KAREN			
TITLE OR INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE			
FILE REFERENCE: 1581.0790001			
CURRENT APPLICATION NUMBER: US/10/320,800			
CURRENT FILING DATE: 2002-12-17			
PRIOR APPLICATION NUMBER: PCT/GB99/03626			
PRIOR FILING DATE: 1999-11-02			
NUMBER OF SEQ ID NOS: 75			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 69			
LENGTH: 939			
TYPE: DNA			
ORGANISM: Neisseria meningitidis			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (1)..(939)			
OTHER INFORMATION: NMB0652 maf A			
US-10-320-800-69			

## Alignment Scores:

Pred. No.: 2,01e-11 Length: 939  
Score: 21.00 Matches: 21  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.56% Indels: 0  
DB: 17 Gaps: 0

US-10-617-835-4 (1-320) x US-10-320-800-69 (1-939)

Qy 214 H1sleuTYAsnAlaGluThrleuLysAlaGlnThrLysLeuGluTYrPheAlaValAsp 233

Db 628 CACCTCTGACAGCGCGAAGCCTTAAGCCCAACCAAGCTCGATTTTCGCCCTGAC 687

Qy 234 Arg 234  
Db 688 CGC 690

## RESULT 6

US-10-027-632-252306/C  
Sequence 252306, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 252306

LENGTH: 2505

TYPE: DNA

ORGANISM: Human

US-10-027-632-252306

## Alignment Scores:

Pred. No.: 12.1 Length: 2505  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.12% Indels: 0  
DB: 13 Gaps: 0

US-10-617-835-4 (1-320) x US-10-027-632-252306 (1-2505)

Qy 4 ArgLeuLeuIleProIleLeuPheSerVal 13

Db 571 AGACTCTTGATACCGATTTTATTTTCGGTG 542

## RESULT 7

US-10-027-632-252306/C  
Sequence 252306, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

TITLE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 252306

LENGTH: 2505

TYPE: DNA

ORGANISM: Human

US-10-027-632-252306

## Alignment Scores:

Pred. No.: 12.1 Length: 2505  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.12% Indels: 0  
DB: 17 Gaps: 0

US-10-617-835-4 (1-320) x US-10-027-632-252306 (1-2505)

Qy 4 ArgLeuLeuIleProIleLeuPheSerVal 13

Db 571 AGACTCTTGATACCGATTTTATTTTCGGTG 542

## RESULT 8

US-09-764-891-9294  
Sequence 9294, Application US/09764891  
Publication No. US20030077808A1  
GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC006

CURRENT APPLICATION NUMBER: US/09/764,891

PRIOR FILING DATE: 2001-01-17

PRIOR APPLICATION data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 10231

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9294

LENGTH: 24173

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-891-9294

## Alignment Scores:

Pred. No.: 94.4 Length: 24173  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.12% Indels: 0  
DB: 10 Gaps: 0

US-10-617-835-4 (1-320) x US-09-764-891-9294 (1-24173)

Qy 140 GlySerGlySerArgSerSerLeuGlyLeu 149

Db 19606 GGCTCAGGCTCCAGTCTCCTCGGCGCTT 19635

RESULT 9  
US-10-091-572-855  
Sequence 855, Application US/10091572  
Publication No. US20030054373A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA118C1  
CURRENT FILING DATE: US/10/091,572  
PRIOR APPLICATION NUMBER: 2002-03-07  
PRIOR FILING DATE: 09/764,850  
PRIOR APPLICATION NUMBER: 2001-01-17  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/218,290  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/225,757  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/226,868  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/216,647  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/225,267  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/216,880  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: 60/225,270  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/251,869  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/235,834  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/234,274  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/234,223  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: 60/228,924  
PRIOR FILING DATE: 2000-08-30  
PRIOR APPLICATION NUMBER: 60/224,518  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,369  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/224,519  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,964  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/241,809  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/249,299  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/236,327  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/241,785  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/244,617  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 60/225,268  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/236,366

PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/251,856  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/251,868  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/229,344  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/234,997  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: 60/229,343  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,345  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,287  
PRIOR FILING DATE: 2000-09-01  
PRIOR APPLICATION NUMBER: 60/229,513  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/231,413  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: 60/229,509  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: 60/236,367  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/237,039  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,038  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/236,370  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: 60/236,802  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,037  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/237,040  
PRIOR FILING DATE: 2000-10-02  
PRIOR APPLICATION NUMBER: 60/240,960  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/239,935  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/239,937  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 60/241,787  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/246,474  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/246,532  
PRIOR FILING DATE: 2000-11-08  
PRIOR APPLICATION NUMBER: 60/249,216  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,210  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/226,681  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225,759  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/225,213  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/227,182  
PRIOR FILING DATE: 2000-08-22  
PRIOR APPLICATION NUMBER: 60/225,214  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/235,836  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/230,438  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/215,135  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: 60/225,266  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/249,218  
PRIOR FILING DATE: 2000-11-17  
PRIOR APPLICATION NUMBER: 60/249,208  
PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: 60/249,213  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,212  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,207  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,245  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,244  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,217  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,211  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,215  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,264  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,214  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/249,297  
 PRIOR FILING DATE: 2000-11-17  
 PRIOR APPLICATION NUMBER: 60/232,400  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/231,242  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,081  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/232,080  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,414  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/231,244  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: 60/233,064  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/233,063  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,397  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,399  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/232,401  
 PRIOR FILING DATE: 2000-09-14  
 PRIOR APPLICATION NUMBER: 60/241,808  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,826  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,786  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/241,221  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/246,475  
 PRIOR FILING DATE: 2000-11-08  
 PRIOR APPLICATION NUMBER: 60/231,243  
 PRIOR FILING DATE: 2000-09-08

Alignment Scores:  
 Pred. No.: 94.4 Length: 24173  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.12% Indels: 0  
 DB: 14 Gaps: 0

US-10-617-835-4 (1-320) x US-10-091-572-855 (1-24173)

QY 140 GlySerGlySerArgSerLeuGlyLeu 149  
 DB 19606 GGCTCAGGCTCCAGGTCTCTCCCTGGGCTT 19635  
 RESULT 10  
 US-10-437-963-48432/c

Sequence 48432, Application US/10437963  
 Publication No. US20040123343A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 APPLICANT: Wu, Wei  
 APPLICANT: Boukharov, Andrey A.  
 APPLICANT: Bardazuk, Brad  
 APPLICANT: Li, Ping  
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 FILE REFERENCE: 38-21(53221)B  
 CURRENT APPLICATION NUMBER: US/10/437,963  
 CURRENT FILING DATE: 2003-05-14  
 NUMBER OF SEQ ID NOS: 204966  
 SEQ ID NO 48432  
 LENGTH: 420  
 TYPE: DNA  
 ORGANISM: Oryza sativa  
 FEATURE:  
 OTHER INFORMATION: Clone ID: PAT\_MRT4530\_51107C.1  
 US-10-437-963-48432

Alignment Scores:  
 Pred. No.: 26 Length: 420  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.81% Indels: 0  
 DB: 19 Gaps: 0

US-10-617-835-4 (1-320) x US-10-437-963-48432 (1-420)

QY 41 ValAlaAlaSerAlaArgAlaAlaVal 49  
 DB 401 GTGGCAGCTCGGCTAGGGCGGCGGTG 375

RESULT 11  
 US-10-425-115-122322/c  
 Sequence 122322, Application US/10425115  
 Publication No. US20040214272A1  
 GENERAL INFORMATION:  
 APPLICANT: La Rosa, Thomas J.  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 FILE REFERENCE: 38-21(53222)B  
 CURRENT APPLICATION NUMBER: US/10/425,115  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 369326  
 SEQ ID NO 122322  
 LENGTH: 463  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURE:  
 OTHER INFORMATION: Clone ID: MRT4577\_43041C.1  
 US-10-425-115-122322

Alignment Scores:  
 Pred. No.: 28.4 Length: 463  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.81% Indels: 0  
 DB: 20 Gaps: 0

US-10-617-835-4 (1-320) x US-10-425-115-122322 (1-463)

QY 40 LeuValAlaAlaSerAlaArgAlaAla 48

Db 442 CTTGTTGCTGCACGTCGCGCCGCG 416

## RESULT 12

US-10-425-115-137587  
 ; Sequence 137587, Application US/10425115  
 ; Publication No. US20040214272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 ; FILE REFERENCE: 38-21(53222)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,115  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 369326  
 ; SEQ ID NO 137587  
 ; LENGTH: 530  
 ; TYPE: DNA  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(530)  
 ; OTHER INFORMATION: unsure at all n locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: MRT4577\_56958C.1  
 ; US-10-425-115-137587

## Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Query Match:	Indels:	Gaps:
32.1	530	9.00	9	100.00%	100.00%	0	2.81%	0	20

US-10-617-835-4 (1-320) x US-10-425-115-137587 (1-530)

Oy 158 ArgAngluThrlleuThrAspPro 166

Db 239 AGGAACGAACCTCACCACGAATCC 265

## RESULT 13

US-09-823-245A-481/c  
 ; Sequence 481, Application US/09823245A  
 ; Publication No. US20020039760A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wong, Gordon G.  
 ; APPLICANT: Clark, Hilary  
 ; APPLICANT: Fechtel, Kim  
 ; APPLICANT: Agostino, Michael J.  
 ; APPLICANT: Howes, Steven H.  
 ; APPLICANT: Resnick, Richard J.  
 ; APPLICANT: Gulukota, Kamalakara  
 ; APPLICANT: Graham, James R.  
 ; APPLICANT: Genetics Institute, Inc.  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
 ; FILE REFERENCE: GIN 6401  
 ; CURRENT APPLICATION NUMBER: US/09/823,245A  
 ; CURRENT FILING DATE: 2001-03-29  
 ; PRIOR APPLICATION NUMBER: 60/194,941  
 ; PRIOR FILING DATE: 2000-04-06  
 ; NUMBER OF SEQ ID NOS: 631  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 481  
 ; LENGTH: 2849  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-823-245A-481

## Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Query Match:	Indels:	Gaps:
147	2849	9.00	9	100.00%	100.00%	0	2.81%	0	9

US-10-617-835-4 (1-320) x US-09-823-245A-481 (1-2849)

Oy 140 GlySerGlySerArgSerSerLeuGly 148

Db 1123 GGCTTGAGTACGCTCTCCTCGGA 1097

## RESULT 14

US-10-311-455-122/c  
 ; Sequence 122, Application US/10311455  
 ; Publication No. US20030143606A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OLEK, Alexander  
 ; APPLICANT: PIEPENBROCK, Christian  
 ; APPLICANT: BERLIN, Kurt  
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
 ; TITLE OF INVENTION: Cytosine methylation  
 ; FILE REFERENCE: 5013.1014  
 ; CURRENT APPLICATION NUMBER: US/10/311,455  
 ; CURRENT FILING DATE: 2002-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
 ; PRIOR FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: DE 10032529.7  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: DE 10043826.1  
 ; PRIOR FILING DATE: 2000-09-01  
 ; NUMBER OF SEQ ID NOS: 2424  
 ; SEQ ID NO 122  
 ; LENGTH: 12705  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: 5677..5678  
 ; OTHER INFORMATION: n is a or g or c or t  
 ; US-10-311-455-122

## Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Query Match:	Indels:	Gaps:
572	12705	9.00	9	100.00%	100.00%	0	2.81%	0	15

US-10-617-835-4 (1-320) x US-10-311-455-122 (1-12705)

Oy 239 LeuLeuIleLysProLysThrAsnAla 247

Db 286 TTACTAATAAACCAGAAACAAACGCC 260

## RESULT 15

US-10-311-455-386/c  
 ; Sequence 386, Application US/10311455  
 ; Publication No. US20030143606A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OLEK, Alexander  
 ; APPLICANT: PIEPENBROCK, Christian  
 ; APPLICANT: BERLIN, Kurt  
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ  
 ; TITLE OF INVENTION: Cytosine methylation  
 ; FILE REFERENCE: 5013.1014  
 ; CURRENT APPLICATION NUMBER: US/10/311,455  
 ; CURRENT FILING DATE: 2002-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537



```

; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 388
; LENGTH: 16720
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-388

```

```

Alignment Scores:
Pred. No.: 733 Length: 16720
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.81% Indels: 0
DB: 15 Gaps: 0

```

US-10-617-835-4 (1-320) x US-10-311-455-388 (1-16720)

```

QY 237 LysLysLeuLeuLleLysProLysThr 245
DB 3994 AAAAAGCTACTATATTAAACCAAACT 3968

```

Search completed: August 19, 2005, 03:44:32  
Job time : 684 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: August 19, 2005, 00:57:16 ; Search time 3360 Seconds  
(without alignments)  
3625.169 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 320

Sequence: 1 MRALLPILPILSVFILSACG.....SHEGYGSDENVRHROGQP 320

## Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US10617835/rnat.18082005.115731.8758/app.query.fasta\_1.519  
-DB=EST -OFMT=fasta -SUPFIX=p2noligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=psio  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10617835 -CGCN 1.1.5180 -rnat.18082005.115731.8758 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESUBSTR -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEX=7

## Database :

EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_hic:  
4: gb\_est3:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_gsa1:  
9: gb\_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	3.1	320	5 BX554222	BX554222 BX554222
2	10	3.1	491	1 AI332340	AI332340 GQ08A04.x
3	10	3.1	528	1 AM005450	AM005450 w286f05.x
4	10	3.1	576	6 CD675737	CD675737 f827c12.y
5	10	3.1	581	1 AI140766	AI140766 q850g02.x
6	9	2.8	254	1 AV326428	AV326428 AV326428
7	9	2.8	254	2 BB557704	BB557704 BB557704
8	9	2.8	404	2 BE092500	BE092500 IL2-BT073
9	9	2.8	441	2 BE434548	BE434548 EST405626

10	9	2.8	455	7 CN493609	CN493609 MdFw2007K
11	9	2.8	464	5 BX091662	BX091662 BX091662
12	9	2.8	479	2 AM037625	AM037625 EST279083
13	9	2.8	485	7 CN284676	CN284676 170005315
14	9	2.8	486	1 AI897283	AI897283 EST266642
15	9	2.8	518	1 AU282041	AU282041 AU282041
16	9	2.8	525	5 BQ397697	BQ397697 NISC_m001
17	9	2.8	528	1 AI486883	AI486883 EST25205
18	9	2.8	530	8 A0663550	A0663550 HS_5476.A
19	9	2.8	543	1 AL779961	AL779961 AL779961
20	9	2.8	554	6 CB878784	CB878784 OSTF087H7
21	9	2.8	572	6 CD290097	CD290097 SLRPU538.
22	9	2.8	575	6 CD319107	CD319107 SLRPU538.
23	9	2.8	595	5 BQ391883	BQ391883 NISC_mq21
24	9	2.8	621	5 BQ392079	BQ392079 NISC_mq22
25	9	2.8	629	4 BG721864	BG721864 602694479
26	9	2.8	625	5 BQ391914	BQ391914 NISC_mq21
27	9	2.8	630	5 CM505798	CM505798 OP_Ba000
28	9	2.8	640	4 BG129256	BG129256 EST744902
29	9	2.8	654	8 AZ572177	AZ572177 302PVB10
30	9	2.8	664	6 CA148027	CA148027 SCEZR2101
31	9	2.8	676	9 BX970295	BX970295 Reverse 8
32	9	2.8	687	9 CE516457	CE516457 C19T-988-
33	9	2.8	702	6 CA147234	CA147234 SCCCR21CO
34	9	2.8	711	8 B2395507	B2395507 E1NCK877R
35	9	2.8	747	9 CC923197	CC923197 C068m12ba
36	9	2.8	769	9 CR215897	CR215897 Forward 8
37	9	2.8	784	5 BX316817	BX316817 BX316817
38	9	2.8	793	6 CB621811	CB621811 OSI18A07P
39	9	2.8	805	6 CB619684	CB619684 OSI18A04G
40	9	2.8	814	7 CF220972	CF220972 AGENCOURT
41	9	2.8	818	7 CK288538	CK288538 EST751260
42	9	2.8	829	9 CC518011	CC518011 CH240_364
43	9	2.8	851	7 CK298583	CK298583 EST761297
44	9	2.8	872	9 CL795658	CL795658 OR_CBA000
45	9	2.8	901	9 CG216003	CG216003 CG2AR86TH

## ALIGNMENTS

RESULT 1  
BX554222  
LOCUS  
DERIVATION  
BX554222 320 bp mRNA linear EST 10-OCT-2003  
morisitans morisitans morisitans adult infected gut Glosina  
morisitans morisitans CDNA clone Tse14D05.q1c, mRNA sequence.

ACCESSION  
BX554222  
VERSION  
BX554222.1 GI:33378317

KEYWORDS  
SOURCE  
ORGANISM  
Glosina morisitans morisitans  
Glosina morisitans morisitans  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Hippoboscidae; Glosinidae; Glosina.

REFERENCE  
AUTHORS  
Lehane, M.J., Aksoy, S., Gibson, W., Keshornou, A., Berriman, M.,  
Hamilton, J., Soares, M.B., Bonaldo, M.F., Lehane, S. and Hall, N.  
Adult midgut expressed sequence tags from the tsetse fly Glosina  
morisitans morisitans and expression analysis of putative immune  
response genes

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
Genome Biol. 4 (10), R63 (2003)  
22881942  
14519198

COMMENT  
Contact: Hall N  
Pathogen Sequencing Unit  
The Sanger Institute The Wellcome Trust Genome Campus  
Hinxton, Cambridge, CB10 1SA, UK  
Request for clones, please contact: Mike Lehane  
Prof. M.J. Lehane  
School of Biological Sciences,  
University of Wales,  
Bangor LL57 2UW

All clones with suffix q1c are reverse primer reads starting at 5' end of the cDNA all pic reads are from

the 3' end.

FEATURES  
SOURCE

Location/Qualifiers  
1..320  
/organism="Glossina morsitans morsitans"  
/mol\_type="mRNA"  
/sub\_species="morsitans"  
/db\_xref="taxon:37546"  
/clone="Tse14p05.q1c"  
/tissue\_type="adult infected gut"  
/clone\_lib="Glossina morsitans morsitans adult infected gut"  
/note="country: zimbabwe; EST from adult gut infected with T.brucei"

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
10.00	10.00	320	10	0	0	0	0
Percent Similarity:	100.00%						
Best Local Similarity:	100.00%						
Query Match:	3.12%						

US-10-617-835-4 (1-320) x BX554222 (1-320)

Oy 237 LysLysLeuLeuLeuLeuProLysProLysThrAsn 246

Db 113 AAAAAATTGTTATTAACCGAAAACTAAC 142

RESULT 2  
AI332340 491 bp mRNA linear EST 28-DEC-1998  
LOCUS 3' mRNA sequence.  
DEFINITION 3' mRNA sequence.  
ACCESSION AI332340  
VERSION AI332340.1 GI:4068899  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 491)

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LINT; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 460.

Location/Qualifiers

1..491

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1931866"

/tissue\_type="pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab\_host="DH10B"

/clone\_lib="Soares NHMPu\_S1"

/note="Organ: mixed (see below); Vector: pT7T3D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NbHM, pregnant uterus

NbHPU, and fetal heart NbH19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools

consisted of I.M.A.G.E. clones 260232-265223,

340488-345479, and 484488-489479."

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
28.1	10.00	491	10	0	0	0	0
Percent Similarity:	100.00%						
Best Local Similarity:	100.00%						
Query Match:	3.12%						

US-10-617-835-4 (1-320) x AI332340 (1-491)

Oy 140 GlySerGlySerArgSerSerLeuGlyLeu 149

Db 470 GGCTCAGGCTCAGCTCCTCCTCGGCTT 441

RESULT 3

AM005450 528 bp mRNA linear EST 20-OCT-2000

LOCUS w286f05.x1 NCI\_CGAP\_Brn25 Homo sapiens cDNA clone IMAGE:2565729 3',

DEFINITION mRNA sequence.

ACCESSION AM005450

VERSION AM005450.1 GI:5854228

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 528)

NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute / National Institute of Neurological

Disorders and Stroke, Brain Tumor Genome Anatomy Project

(CGAP/BTGA), Tumor Gene Index

Unpublished (1998)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lemmon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LINT at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 957 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 466.

Location/Qualifiers

1..528

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2565729"

/tissue\_type="anaplastic oligodendroglioma"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP\_Brn25"

/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a

modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGGCGCCGATGATGTTTATTTTATTTTATTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

Library is normalized, and was constructed by Bento

Soares and M. Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:
30	10.00	528	10	0	0
Percent Similarity:	100.00%				
Best Local Similarity:	100.00%				

Query Match: 3.12% Indels: 0  
DB: 2 Gaps: 0  
US-10-617-835-4 (1-320) x AW005450 (1-528)

Qy 140 Glycerglyserargserleuclyleu 149  
DB 474 GGCTCAGGCTCCAGGCTCCTCGGCGCTT 445

RESULT 4  
LOCUS CD675737 576 bp mRNA linear EST 24-JUN-2003  
DEFINITION f827c12.y1 Human lens cDNA (Normalized): fs Homo sapiens cDNA clone  
f827c12.5, mRNA sequence.  
ACCESSION CD675737  
VERSION CD675737.1 GI:32177468  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 576)  
Wistow,G., Bernstein,S.L., Wyatt,M.K., Behal,A., Touchman,J.W.,  
Bouffard,G., Smith,D. and Peterson,K.  
Expressed sequence tag analysis of adult human lens for the NEIBank  
Project: over 2000 non-redundant transcripts, novel genes and  
splice variants  
Mol. Vis. 8 (4), 171-184 (2002)  
JOURNAL 22103463  
MEDLINE 12107413  
PUBMED

COMMENT Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/31, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: gwaem@nei.nih.gov  
Plate: 27 row: C column: 12  
Seq primer: M13RP1 reverse primer (ABI).  
Location/Qualifiers  
1..576  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="f827c12"  
/issue\_type="Lens"  
/dev\_stage="Adult"  
/lab\_host="EMD110B"  
/clone\_lib="Human Lens cDNA (Normalized): fs"  
/note="Organ: Eye; Vector: pCMVSPORT6; A human lens  
library (by) was normalized by self-subtraction. One  
portion of double stranded plasmid DNA representing the  
library was linearized by NotI. This NotI digested library  
was used as a template for biotinylated DNA synthesis  
using SP6 RNA polymerase. Another portion of the double  
stranded plasmid library was converted to single-stranded  
circles in vitro using Gene II and Exonuclease III (Life  
Technologies). Single-stranded DNA (1 mg) was hybridized  
(Cot 500) with 41 mg of Bio-RNA and vector blocking  
oligonucleotides. The hybridized Bio-RNA/ss-circles were  
removed by streptavidin:phenol extraction. EST analysis  
was performed on the library at the NIH Intramural  
Sequencing Center (NISC)."

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
33.6	10.00	576	10	0	0	0	0
100.00%	100.00%						
Best Local Similarity:	3.12%						
Query Match:	6						
DB:							

US-10-617-835-4 (1-320) x CD675737 (1-576)

Qy 140 Glycerglyserargserleuclyleu 149  
DB 484 GGCTCAGGCTCCAGGCTCCTCGGCGCTT 513

RESULT 5  
LOCUS A1140766 581 bp mRNA linear EST 05-OCT-1998  
DEFINITION q450g02.x1 Soares\_NHMPu\_S1 Homo sapiens cDNA clone IMAGE:1690226  
3', mRNA sequence.  
ACCESSION A1140766  
VERSION A1140766.1 GI:3648223  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 581)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
Email: cga@bcr-remail.nih.gov  
This clone is available royalty-free through LUNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 782 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 444.  
Location/Qualifiers  
1..581  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1690226"  
/issue\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NHMPu\_S1"  
/note="Organ: mixed (see below); Vector: p77T3D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI; Equal amounts of plasmid DNA from three  
NBHPU, and fetal heart NBH19W, were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."

ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
32.9	10.00	581	10	0	0	0	0
100.00%	100.00%						
Best Local Similarity:	3.12%						
Query Match:	1						
DB:							

US-10-617-835-4 (1-320) x A1140766 (1-581)

Qy 140 Glycerglyserargserleuclyleu 149  
DB 478 GGCTCAGGCTCCAGGCTCCTCGGCGCTT 449

RESULT 6  
LOCUS AV326428 254 bp mRNA linear EST 11-NOV-1999  
DEFINITION AV326428 RIKEN full-length enriched, adult male medulla oblongata  
mus musculus cDNA clone 633041A18 3' similar to AF009513 Mus  
musculus hematopoietic lineage switch 2 (HLS2) mRNA, mRNA sequence.



/note="Vector: pBluescriptSKmCudaapc; Site\_1: EcoRI; Site\_2: XhoI; Fruit were harvested at the Breaker stage (first sign of lycopene accumulation on the blossom end of

the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

## ORIGIN

## Alignment Scores:

Pred. No.:	251	Length:	441
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.81%	Indels:	0
DB:	2	Gaps:	0

US-10-617-835-4 (1-320) x BE434548 (1-441)

Oy 76 SerLeuThrg1Yg1YArgrTYrSer1le 84

Db 219 AGCTTGACCGGCGCGGATCTCCATC 193

## RESULT 10

CN493609

LOCUS Mdfw2007k07.y1 Mdfw Malus x domestica cDNA linear EST 24-MAY-2004

DEFINITION similar to SW:TL16.ARAH 022773 TRYKOID LUMENAL 16.5 KDA PROTEIN, CHLOROPLAST PRECURSOR, mRNA sequence.

ACCESSION CN493609

VERSION CN493609.1 GI:46613899

KEYWORDS EST.

SOURCE Malus x domestica (cultivated apple)

## REFERENCE

AUTHORS

Alwinckle, H., Mainio, M., Carroll, N., Goldbrogh, P., Ovis, K., Clifton, S., Pape, D., Marra, M., Hillier, J., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Ronko, I., Tsagaris, R., Kennedy, S., Waterston, R. and Wilson, R.  
Apple Functional Genomics grant - NSF 0321702  
Unpublished (2004)

## JOURNAL

COMMENT Contact: Schuyler S. Korban  
Apple Functional Genomics grant - NSF 0321702  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: esat@wustl.edu  
Library materials provided by: Schuyler S. Korban Library  
constructed by: A. Hernandez / K. Gasic Library sequenced by:  
Washington University Genome Sequencing Center  
Washu EST name: aaf25f04.y1.

location/Qualifiers

1..455

/organism="Malus x domestica"

/mol\_type="mRNA"

/db\_xref="taxon:3750"

/clone="Mdfw2007k07"

/lab\_host="DH10B ampicillin resistant"

/clone\_lib="Mdfw"

/note="Vector: DH10B ampicillin resistant, Site 1: NotI, Site 2: EcoRI; Total RNA was extracted separately from each stage (bud, balloon, open and after pollination), using the 'pine tree' method. Poly(A)+mRNA was isolated twice from total RNA from each stage using the Oligotex Direct mRNA kit (Qiagen). mRNA was reverse transcribed into double stranded cDNA using a modified oligo18(dT) primer with an identifying tag sequence (see table below). cDNAs from different stages were pooled in equal amounts before adaptor ligation. Tag identification when sequencing from 5' end: Stage 1 (bud) insert 18(A)TCGGA; Stage 2 (balloon) insert 18(A)TCGA; Stage 3 (open) insert 18(A)TCGT; Stage 4 (after pollination) insert 18(A)TCGT. Tag identification when sequencing from 3' end: Stage 1

## ORIGIN

## Alignment Scores:

Pred. No.:	259	Length:	455
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.81%	Indels:	0
DB:	7	Gaps:	0

US-10-617-835-4 (1-320) x CN493609 (1-455)

Oy 1 MeArgrAlaargLeuLeuIleProIle 9

Db 389 ATGAGACCAAGACTATTAATTCACATA 415

## RESULT 11

BX091662/c

LOCUS BX091662 Soares NFL T GBC S1 Homo sapiens cDNA clone

DEFINITION IMAGE98E075157, IMAGE:2092014, mRNA sequence.

ACCESSION BX091662

VERSION BX091662.1 GI:27824920

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 464)

AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Patsch, E., Peters, M., Radloff, U., Schneider, D. and Korn, B.

TITLE Human Unigeneset - RZPD3

JOURNAL Human Unigeneset - RZPD3

COMMENT Unpublished (2003)

Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 360, D-69120 Heidelberg, Germany  
RZPD; IMAGE98E075157.  
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No.972)  
http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany



Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r, Primer sequence: TTTCCACAGCAACAGCTATGAC.

## FEATURES

## SOURCE

1.464  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE98B075157 ; IMAGE:2092014"  
/lab\_host="DH10B"  
/clone\_lib="Soares\_NFL\_T\_GBC\_S1"  
/note="Organ: pooled; Vector: pTV73D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
Equal amounts of plasmid DNA from three normalized  
libraries (fetal lung NBHL19W, testis NHT, and B-cell  
NCI-GAP GCBI) were mixed and as circles were made in  
vitro. Following HAP purification, this DNA was used as  
tracer in a subtractive hybridization reaction. The driver  
was PCR-amplified cDNAs from pools of 5,000 clones made  
from the same 3 libraries. The pools consisted of  
I.M.A.G.E. clones 297480-302087, 682632-687239,  
726408-728711, and 729096-731399. Subtraction by Bento  
Soares and M. Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.:	264	Length:	464
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.81%	Indels:	0
DB:	5	Gaps:	0

US-10-617-835-4 (1-320) x BX091662 (1-464)

Qy 9 |leleupheservalelphelleusar 17

Db 232 ATTCTTCTCTGTGTCTACTACTGTC 206

## RESULT 12

AW037625 479 bp mRNA linear EST 18-MAY-2001

## LOCUS

DEFINITION EST179083 tomato mixed elicitor, B71 Lycopersicon esculentum cDNA

clone CLERT309, mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## ORGANISM

## SOURCE

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## SOURCE

## LOCATION/Qualifiers

## 1.479

## /organism="Lycopersicon esculentum"

## /mol\_type="mRNA"

## /cultivar="Rio Grande Flor"

## 5 prime sequence.

## Location/Qualifiers

## 1.479

## /organism="Lycopersicon esculentum"

## /mol\_type="mRNA"

/db\_xref="taxon:4081"  
/clone="CLERT309"  
/tissue\_type="leaf"  
/dev\_stage="4-6 week old plants"  
/lab\_host="XLI-Blue MRF"  
/clone\_lib="tomato mixed elicitor, B71"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
XhoI; CLERT - inoculated with a variety of disease response  
elicitors. Plants exposed to 2,6 dichloroisocitric  
acid, BTH, jasmonic acid, ethylene, fenchon, EIX,  
okadaic acid, or systemin prior to tissue harvest. EcoRI  
site was destroyed during cloning."

## ORIGIN

Pred. No.:	271	Length:	479
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.81%	Indels:	0
DB:	2	Gaps:	0

US-10-617-835-4 (1-320) x AW037625 (1-479)

## Qy 76 Serleutnrglyglyargtyrserile 84

Db 152 AGCTTGACCGGCGCGACTTCATC 126

## RESULT 13

CN284676 485 bp mRNA linear EST 16-MAY-2004

## LOCUS

DEFINITION 17000531558899 GRN\_ES Homo sapiens cDNA 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## ORGANISM

## SOURCE

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## SOURCE

## LOCATION/Qualifiers

## 1.485

## /organism="Homo sapiens"

## /mol\_type="mRNA"

## /db\_xref="taxon:9606"

## /clone\_lib="GRN\_ES"

/note="Oligo dT primed, full-length enriched cDNA library  
from undifferentiated hbs cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## H9"

## /clone\_lib="GRN\_ES"

/note="Oligo dT primed, full-length enriched cDNA library  
from undifferentiated hbs cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## H9"

## /clone\_lib="GRN\_ES"

/note="Oligo dT primed, full-length enriched cDNA library  
from undifferentiated hbs cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## H9"

## /clone\_lib="GRN\_ES"

/note="Oligo dT primed, full-length enriched cDNA library  
from undifferentiated hbs cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## H9"

## /clone\_lib="GRN\_ES"

/note="Oligo dT primed, full-length enriched cDNA library  
from undifferentiated hbs cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

## H9"

## /clone\_lib="GRN\_ES"

/note="Oligo dT primed, full-length enriched cDNA library  
from undifferentiated hbs cell lines H1 (p32), H7 (p29),  
and H9 (p26) maintained in feeder-free conditions"

DB: 7 Gaps: 0

US-10-617-835-4 (1-320) x CN284676 (1-485)

OY 140 G|ySerg|ySerg|ySerg|Ser|en|g|y 148  
 |||||  
 DB 381 GGCTCTGGGTACGCTCCTCCCTGGGA 355

RESULT 14  
 A1897283/C 486 bp mRNA linear EST 18-MAY-2001

LOCUS EST26642 tomato ovary, TAMU Lycopersicon esculentum cDNA clone

DEFINITION EST26642, mRNA sequence.

ACCESSION A1897283

VERSION A1897283.1 GI:5603101

KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 486)  
 Alcalá, J., Vrebalov, J., White, R., Matern, A. L., Vision, T.,  
 Holt, I. E., Liang, F., Upton, J., Romning, C. M., Craven, M. B.,  
 Fujii, C. Y., Bowman, C. L., Niernan, W., Fraser, C. M., Venter, J. C.,  
 Martin, G. B., Tanksley, S. D. and Giovannoni, J.  
 Generation of ESTs from tomato carpel tissue  
 Unpublished (1999)

TITLE JOURNAL

COMMENT Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

FEATURES  
 Location/Qualifiers  
 1..486  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLD26N1"  
 /tissue\_type="carpel"  
 /dev\_stage="5 days pre-anthesis to 5 days post-anthesis"  
 /lab\_host="X11-Blue MRP"  
 /clone\_lib="tomato ovary, TAMU"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; cLED - Tomato Carpel EST Library. OligodT-primed and  
 directionally cloned cDNA in vector Lambda ZAP II with 5'  
 and 3' ends located at the EcoRI and XhoI sites,  
 respectively."

ORIGIN

Alignment Scores:

Pred. No.:	275	Length:	486
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.81%	Indels:	0
DB:	1	Gaps:	0

US-10-617-835-4 (1-320) x A1897283 (1-486)

OY 76 Ser|euthr|g|y|g|y|A|g|y|S|er|le 84  
 |||||  
 DB 266 AGCTTGACCGGCGCGCATCTCCATC 240

RESULT 15  
 AU282041 518 bp mRNA linear EST 03-DEC-2002

LOCUS AU282041 Molgula tectiformis library Molgula tectiformis cDNA clone

DEFINITION MT19B2G11T 3', mRNA sequence.

ACCESSION AU282041

VERSION AU282041.1 GI:26021539

KEYWORDS EST.

SOURCE Molgula tectiformis

ORGANISM Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 Stolidobranchia; Molguliidae; Molgula.  
 1 (bases 1 to 518)  
 Takada, N., Murakami, S. D., Swalla, B. J. and Satch, N.  
 EST analysis of tailless ascidian M. tectiformis  
 Unpublished (2002)

TITLE JOURNAL

COMMENT Contact: Seiko D. Murakami  
 Graduate School of Science, Department of Zoology  
 Kyoto University  
 Sakyo-ku, Kyoto 606-8502, Japan  
 Tel: 81-75-753-4095  
 Fax: 81-75-705-1113  
 Email: [seiko@ascidian.zool.kyoto-u.ac.jp](mailto:seiko@ascidian.zool.kyoto-u.ac.jp).

FEATURES  
 Location/Qualifiers  
 1..518  
 /organism="Molgula tectiformis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:30286"  
 /clone="MT19B2G11T"  
 /clone\_lib="Molgula tectiformis library"

ORIGIN

Alignment Scores:

Pred. No.:	292	Length:	518
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.81%	Indels:	0
DB:	1	Gaps:	0

US-10-617-835-4 (1-320) x AU282041 (1-518)

OY 169 Thr|A|p|he|u|S|er|H|s|le|u|A|G|n 177  
 |||||  
 DB 391 ACTGCTTCTCTGTCATCTGTCAG 417

Search completed: August 19, 2005, 02:17:54  
 Job time : 3378 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 19, 2005, 01:21:31, Search time 533 Seconds

(without alignments)  
3554.066 Million cell updates/sec

Title: US-10-617-835-4

Perfect score: 320  
Sequence: 1 MRARLLRILFSLVFLSACG.....SHEGIVGSDAVRHRCGQP 320

Scoring table:

OLIGO	
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp  
-Q=/cgn2.1/USPRO.spool.p/US10617835/runat.18082005.115729.8741/app\_query.fasta.1.519  
-DB=N Geneseq.16Dec04 -QFWT=fastcap -SUFFIX=p2nol1go.rng -MINMATCH=0.1  
-LOOFCLO=0 -LOOPEXT=0 -UNITS=bites -SPART=1 -END=1 -MATRIX=ol1go  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1  
-ALLEN=15 -MODE=LOCAL -OUTFWT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10617835\_@CGN\_1.1.708\_@runat.18082005.115729.8741  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq.16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	320	100.0	3287	2 AAT70386	AAT70386 Neisseria
2	282	88.1	960	10 ABZ703917	ABZ703917 N. gonorr
3	282	88.1	963	2 AAZ12008	AAZ12008 Neisseria
4	282	88.1	963	3 AA254614	AA254614 Neisseria
5	282	88.1	963	3 AA253688	AA253688 Neisseria

6	282	88.1	963	3 AAAB1297	AAAB1297 N. gonorr
7	282	88.1	963	3 AAF21582	AAF21582 N. gonorr
8	257	80.3	789	10 ABZ40799	ABZ40799 N. gonorr
9	207	64.7	963	2 AAZ12006	AAZ12006 Neisseria
10	207	64.7	963	3 AA253689	AA253689 Neisseria
11	207	64.7	963	3 AA254613	AA254613 Neisseria
12	207	64.7	963	3 AAAB1296	AAAB1296 N. mening
13	207	64.7	963	3 AAF21581	AAF21581 N. mening
14	194	60.6	92834	3 AAAB1473	AAAB1473 N. mening
15	194	60.6	110000	3 AAAB189_7	Continuation (8 of
16	194	60.6	172325	3 AAF21613	AAF21613 Neisseria
17	179	55.9	963	2 AAZ12007	AAZ12007 Neisseria
18	179	55.9	963	3 AA254615	AA254615 Neisseria
19	179	55.9	963	3 AA253690	AA253690 Neisseria
20	179	55.9	963	3 AAAB1298	AAAB1298 N. mening
21	179	55.9	963	3 AAF21583	AAF21583 N. mening
22	69	21.6	640	2 AAZ12005	AAZ12005 Neisseria
23	69	21.6	640	3 AAAB1330	AAAB1330 N. mening
24	55	17.2	399	3 AAAB2165	AAAB2165 N. mening
25	25	6.6	882	10 ABZ39704	ABZ39704 N. gonorr
26	21	6.6	935	10 ABZ39696	ABZ39696 N. gonorr
27	21	6.6	939	10 ABZ41775	ABZ41775 N. gonorr
28	21	6.6	942	2 AAZ12110	AAZ12110 Neisseria
29	21	6.6	942	2 AAZ12109	AAZ12109 Neisseria
30	21	6.6	25509	3 AAAB1467	AAAB1467 N. mening
31	21	6.6	110000	3 AAAB1490_03	Continuation (4 of
32	21	6.6	110000	3 AAAB1490_06	Continuation (7 of
33	21	6.6	110000	3 AAAB1489_0	AAAB1489 N. mening
34	21	6.6	349980	3 AAF21607	AAF21607 Neisseria
35	21	6.6	349980	3 AAF21608	AAF21608 Neisseria
36	17	5.3	591	2 AAAB1359	AAAB1359 N. mening
37	17	5.3	591	3 AAAB1359	AAAB1359 N. mening
38	17	5.3	591	3 AAAB1359	AAAB1359 N. mening
39	10	3.1	24173	4 AAK79926	AAK79926 Human Imm
40	10	3.1	24173	4 AAL06606	AAL06606 Human rep
41	10	3.1	24173	5 AA540703	AA540703 DNA encod
42	10	3.1	24173	11 ADJ09909	ADJ09909 Human pro
43	9	2.8	758	4 AAH03174	AAH03174 Human cDN
44	9	2.8	2849	6 ABK34712	ABK34712 Human cDN
45	9	2.8	12705	6 ABJ32149	ABJ32149 Human Imm

#### ALIGNMENTS

RESULT 1  
AAT70386  
ID AAT70386 standard; DNA; 3287 BP.  
AC AAT70386;  
XX  
AC AAT70386;  
XX  
DT 08-DEC-1997 (first entry)  
XX  
DE Neisseria adhesion proteins.  
XX  
KW Neisseria gonorrhoeae; adhesion; lipoprotein; OrfA; OrfI; OrfB; ss.  
XX  
OS Neisseria gonorrhoeae.  
XX  
FH  
FT Key  
FT -35\_signal  
FT Location/Qualifiers  
FT 63..68  
FT /\*tag= a  
FT 78..83  
FT /\*tag= b  
FT 124..128  
FT RBS  
FT 136..449  
FT /\*tag= c  
FT CDS  
FT 136..449  
FT /\*tag= d  
FT /label= OrfA  
FT 514..519  
FT /\*tag= e  
FT 534..539  
FT /\*tag= f  
FT 571..577  
FT RBS

```

FT      /tag= g
FT      583..1545
FT      /*tag= h
FT      /label= Orf1
FT      1573..1579
FT      /*tag= i
FT      1585..3114
FT      /*tag= j
FT      /label= OrfB
XX      DE19534579-A1.
XX
XX      20-MAR-1997.
XX
XX      18-SEP-1995; 95DE-01034579.
XX
XX      18-SEP-1995; 95DE-01034579.
XX
XX      (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX
XX      WPI; 1997-180942/17.
XX      P-PSDB; AAM18784, AAM18785, AAM18786.
XX
XX      Nucleic acids encoding Neisseria adhesion proteins - for therapeutic and
XX      diagnostic use.
XX      Claim 1; Page 11-13; 20pp; German.
XX
XX      OrfA and OrfB in complexes with the protein PilC are capable of adhering
XX      to human cells. Products obtained from the DNA are useful in medicaments,
XX      diagnostic compns. and vaccines, esp. for treatment of Neisseria
XX      gonorrhoea and N. meningitidis infections
XX
XX      Sequence 3287 BP; 1016 A; 741 C; 752 G; 778 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 0 Length: 3287
Score: 320.00 Matches: 320
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-617-835-4 (1-320) x AAT70386 (1-3287)
QY      1 MetArgAlaArgLeuLeuLeuProIleuRhesSerValPheIleLeuSerAlaCysGly 20
DB      583 ATGGGGGACGGCTGCTGATACCTATCTTTTCGTTTATTTATTCGCCCTCGCGGG 642
QY      21 ThrLeuThrgIyIleProSerHisGlyGlyIyIySArGpheaIaValGluGlnGluLeu 40
DB      643 ACACGTGACAGGTATTCATCGCATGGCGGAGCAAAAGCTTCGCGGTGAAACAAGACTT 702
QY      41 ValAlaAlaSerAlaAlaGAlaAlaValIySArpMetApleuGlnAlaLeuHisGlyArg 60
DB      703 GTGGCCCGCTCTGCCAGAGCTGCCGTTAAAGACATGATTTTACAGGCACTTACAGGAGA 762
QY      61 LysValAlaLeuYrIleAlaThrMeGlyVaspgInglySerGlySerLeuThrGlyGly 80
DB      763 AAAGTTGCATTTGATTCATTCGCAACTATGGGCAACCAAGTTTCAAGGAGTTTACAGGGGGT 822
QY      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyIyIyIleAsnSerProAlaValArgThr 100
DB      823 CGCTACTCCATTTGATGACGCTGATTGCGGGCAATTCATTAACAGGCTCGCGCACCC 882
QY      101 AspYrYrThrYrProArgYrGlyIuThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
DB      883 GATTACACCTATCCGCGTTACGAAACCAACCGCTGTAACCAACATCAGGCGTTTGAACGGGT 942
QY      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
DB      943 TTAACCACTTCTTATCTACACTTAATGCCCCCTGCACTCTCGCGCACCAATCAGACGGT 1002

```

```

QY      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspYrArgAsnGlu 160
DB      1003 AGCGGAAGTAGACACAGCTCTGGCTTAATATTGGCGGGATGGGGGATTAACGAATGAA 1062
QY      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
DB      1063 ACCTTGACGACCAACCCCGCGGACACCTGCTTTTCCCACTTGATACACACGTAATT 1122
QY      181 PheLeuArgGlyIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsn 200
DB      1123 TTCCCTGCGCGCATAGACGTTGTTTCTTCCTCCCAATCCCGATACAGATGTGTTATTAC 1182
QY      201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuYrAsnAlaGluThr 220
DB      1183 ATTCACGTAATTCGAGACGATACGCAACGAAACCGAAATGCACCTATACAAATGCCGAAACA 1242
QY      221 LeuIySAlaGlnThrIySleuGluIyTyrPheAlaValAspArgThrAsnIySleuLeu 240
DB      1243 CTGAAGCCCAAAACAAACCTGGAATATTTCGCAGTAGACAGAACCAATMAAAATGCTC 1302
QY      241 IleYsProIySThrAsnAlaPheGluAlaIatYrIySgluAsnYrAlaLeuTrpMet 260
DB      1303 ATCAAAACCCAAACCAATGCGTTTGAAGCTGCTCTTAAGAAATTAACGACTTGATG 1362
QY      261 GlyProTyrIySValSerIySglyIleIySProThrgIyLeuMetValAspPheSer 280
DB      1363 GGGCCGTATAAAGTAACCAAGAAATCAACCGACGAGAGATTAATGCTGATTTCTCC 1422
QY      281 AspIleArgProTyrGlyIyAsnHisThrGlyAsnSerIaProSerValGluAlaAspAsn 300
DB      1423 GATATCCGGCCATACGGCAATCATACGAGTAATCTCCGCCCATCCGTAGAGGCTGATTAAC 1482
QY      301 SerHisgluGlyTyrGlyIyTyrSerAspGluAlaValArgGlnHisArgGlnGlyIyPro 320
DB      1483 AGTCATGAGGGGTATGATACAGCGATGAGCACTGCGACACATACGACAAAGGCAACCT 1542

RESULT 2
ABZ37917
ID      ABZ37917 standard; DNA; 960 BP.
XX
XX      ABZ37917;
XX
XX      07-MAR-2003 (first entry)
XX
XX      N. gonorrhoeae nucleotide sequence SEQ ID 423.
XX
XX      Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
XX      Neisseria gonorrhoeae.
XX
XX      W0200279243-A2.
XX
XX      10-OCT-2002.
XX
XX      12-FEB-2002; 2002WO-1B002069.
XX
XX      12-FEB-2001; 2001GB-00003424.
XX
XX      (CHIR-) CHIRON SPA.
XX
XX      Fontana MR, Pizza M, Maignani V, Monaci E;
XX      WPI; 2003-058415/05.
XX      DR      P-PSDB; ABP76947.
XX
XX      New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX      PT      medicament for treating or preventing N. gonorrhoeae infection.
XX      PS      Disclosure; Page 220; 815pp; English.
XX
XX      The present invention relates to proteins from Neisseria gonorrhoeae.
XX      CC      Also disclosed are the nucleic acid molecules encoding the proteins and
XX      CC      antibodies that specifically bind to the proteins. The composition

```



US-10-617-835-4 (1-320) x AAZ12008 (1-963)

OY	1	MetAga1aArgLeuLeu1lePro1leLeuPheSerValPhe1leuSer1aAcYsgIy	20
Db	1	ATGCGGGACGGCTGCTGATACCCTATTCTTTTTCAGTTTATTTTATTCGGCTCGGG	60
OY	21	ThrLeuThrGly1leProSerHisglYgIyLysArgPheAlaValgluInglueu	40
Db	61	ACACTGACGAGGATTCCTCGCATGGCGGAGGCAACGGCTTCGGGTGCAACAAGACTT	120
OY	41	ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGln1leuHisglYArg	60
Db	121	GTGGCGGCTTCGCCAGAGCTCGCTGTAAGAATGATGATTTTACAGCATTCACGGACGA	180
OY	61	LysValAlaLeuTyr1leAlaThrMetCglYaspGlnglYSerCglYserLeuThrngIyglY	80
Db	181	AAAGTTGACATTGTACATTGCACACTATGGCGGACCAAGGTTCAAGGAGTTTGA	240
OY	81	ArgTyrSer1leAspAlaLeu1leArgglYglYTyrl1leAsnSerProAlaValArgTTr	100
Db	241	CGCTACTCCATTGATGCACTGATTCGGCGGAATACATAAACGCCCTCGCGCCGACCC	300
OY	101	AspTyrThrTyrProArgTyrCgluThrThrAlaGluThrThsSerGlyLeuThrngIy	120
Db	301	GATTACACTATCCGGGTTACGAAACACCGGTGAACAACATCAGCGGTTTGA	360
OY	121	LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly	140
Db	361	TTTAAACCACTCTTTTATCTACACTAAATGGCCCTCGCACTCGGGACCCCAATACAGGT	420
OY	141	SerGlySerArgSerSerLeuGlyLeuAsn1leGlylYmetCglYAspTyrArgAsnGlu	160
Db	421	AGCGAAGTAGGAGCAGCTGTGGGCTTAAATTTGGGGATGGGGGATTAATCGAAATGAA	480
OY	161	ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe	180
Db	481	ACCTTGACGACCAACCCCGGCGACACTGCTCTTCTTCCCACTTGATGACAGCCGATTT	540
OY	181	PheLeuArgGly1leAspValValSerProAlaAsnAlaAspThrAspValPhe1leAsn	200
Db	541	TTTCGTGGCGGGATGAGCGTTGTCTCTCGCCAAATGCCGATACAGATGTTATTTAAC	600
OY	201	1leAspValPheGlyThr1leArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr	220
Db	601	ATCGACGATTCGGAGACGATACGCACAACAACCGAATCACTATACATATGCCGAACA	660
OY	221	LeuLysAlaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLysLeuLeu	240
Db	661	CTGAAAGCCCAACAAAATCGGATATTTTCGCAGTGACAGAACCAATATAAAAAATTGCTC	720
OY	241	1leLysProLysThrAsnAlaPheGluAlaAlaTyrLysGluAsnTyrAlaLeuTTPMet	260
Db	721	ATCAAAACCCAAAACAATCGTTTGAAAGCTGCTCTATAAGAAAATTACGATTTGTGATG	780
OY	261	GlyProTyrLysValSerLysglY1leLysProThrGlnglYLeuMetValAspPheSer	280
Db	781	GGGCGGCTATAAGTAAAGCAAGGAATCAACACGACGGAAGGATTAATGTCATTTCTCC	840
OY	281	Asp1leArg-ProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAs	300
Db	841	GATATCCA-ACCATACGGCAATCATATCGGGTAACTCCGCCCATTCGTAGAGGCTATTA	899
OY	300	nSerHisgluTyrTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnglYglInP	320
Db	900	CAGTATATGAGGGGTATGATACGACGATGACGATGACGACATACACAAATACACAGGCAACC	959
OY	320	O 320	
Db	960	T 960	

RESULT 4  
AAZ54614  
ID AAZ54614 standard; DNA; 963 BP.

XX	AAZ54614;	
AC		
XX		
DT	15-SEP-2003 (revised)	
DT	21-MAR-2000 (first entry)	
XX		
DE	Neisseria gonorrhoeae ORF 406 partial DNA sequence SEQ ID NO:3111.	
XX		
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;	
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;	
XX	antibacterial; gene therapy; ds.	
XX		
OS	Neisseria gonorrhoeae.	
XX		
PN	W09957280-A2.	
XX		
PD	11-NOV-1999.	
XX		
PF	30-APR-1999; 99WO-US009346.	
XX		
PR	01-MAY-1998; 98US-0083758P.	
PR	31-JUL-1998; 98US-0094869P.	
PR	02-SEP-1998; 98US-0098994P.	
PR	02-SEP-1998; 98US-0099062P.	
PR	09-OCT-1998; 98US-0103749P.	
PR	09-OCT-1998; 98US-0103794P.	
PR	09-OCT-1998; 98US-0103796P.	
XX	25-FEB-1999; 99US-0121528P.	
XX		
PA	(CHIR ) CHIRON CORP.	
PA	(GENO ) INST GENOMIC RES.	
XX		
PI	Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;	
PI	Petersen J, Piza M, Rappoli R, Ratti G, Scalato E, Scarselli M;	
PI	Tetzelin H, Venter JC;	
XX		
DR	WPI: 2000-062150/05.	
DR	P-PSDB; AAY75812.	
XX		
PT	Novel Neisserial polypeptides predicted to be useful antigens for	
PT	vaccines and diagnostics.	
XX		
PS	Example 1; Page 109; 145pp; English.	
XX		
CC	AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941	
CC	represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides	
CC	and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent	
CC	PCR primers used in the exemplification of the present invention. The	
CC	polypeptides, the polynucleotides, antibodies and compositions of the	
CC	invention can be used as vaccines, as diagnostic reagents, and as	
CC	immunogenic compositions. The polypeptides can be used in the manufacture	
CC	of medicaments for treating or preventing infection due to Neisserial	
CC	bacteria (e.g. meningitis and septicemia), to detect the presence of	
CC	Neisseria bacteria, or to raise antibodies. They may also be used to	
CC	screen for agonists or antagonists, which may themselves have use as	
CC	antibacterial agents. The polynucleotides of the invention may also be	
CC	used in gene therapy protocols. (Updated on 15-SEP-2003 to standardise OS	
CC	field)	
XX		
XX	Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;	

US-10-617-835-4 (1-320) X AAZ54614 (1-966)  
QY 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20

```

Db      1 ATGGGGGACGGGCTGCTGATACCTATCTTTTTCAGTTTATTTTATCCGCGCTGGCGG 60
Qy      21 ThrLeuThrGlyIleProSerHisglYglYlyValGPhenAlaValGluGlnGluLeu 40
Db      61 ACACGACAGCAGGTATTCATCGCATGCGGAGGCAACGCTTCGCGGTCCACAAAGAACTT 120
Qy      41 ValAlaIaSerAlaArgAlaAlaValLyValAspMetAspLeuGlnAlaLeuHisglYArg 60
Db      121 GTGGCGGCTTCTGCGAGAGCTCCGTTAAAGACATGAGATTTCAGGCATTACCGGACCA 180
Qy      61 LyValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Db      181 AAAGTTGCACTTATCACTTCACTAATGCGGCGAACCAAGGTTCAAGGAGTTGACAGGGGCT 240
Qy      81 ArgTyrSerIleAspAlaLeuIleArgGlyIuTyrIleAsnSerProAlaValArgThr 100
Db      241 CGCTACTCCATTTGATGACCTGATGCGCGGCAATACATAAACAGCCCTGCGCGCACCC 300
Qy      101 AspTyrThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGly 120
Db      301 GATTACACCTATCCCGGTTACGAAACCAACGCTGAAACCAACATCAAGCGGTTGACGGGT 360
Qy      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db      361 TTAAACCACTTCTTTATCTACACTTAATGCCCCCTGCACTCTCGCGACCCCAATCAGACG 420
Qy      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAnglu 160
Db      421 AGCGGAGAGAGAGAGAGTCTGCGCTTAAATATTGGCGGAGATGGGGGATTTATGAAATGA 480
Qy      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      481 ACCTTGACACCAACCCCGGCAACATGCTCTTCTTCCACCTGGTGACACCGTATTT 540
Qy      181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541 TTCCTGCGCGGACATGAGCTGTTCTCTCTGCGCAATGCCGATACAGATGTTTATTAC 600
Qy      201 IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThr 220
Db      601 ATCGACGATTCGGAACGATACGCAACAGAACCGAAATCACCTATACAAATCCGAAACA 660
Qy      221 LeuValAlaGlnThrIleLeuGluIuTyrPheAlaValAspArgThrAsnIleLeuLeu 240
Db      661 CTGAAAGCCCAACAAACCTGGAATATTTGCGAGTAGACAGAACCAATTAATAAATTGCTC 720
Qy      241 IleLysProLysThrAsnAlaPheGluAlaIaTyrIleGlyAsnTyrAlaLeuTyrPhe 260
Db      721 ATCAAAACCAAAACCAATGCTTGAAGCTGCTATTAAGAAATTAACCATTTGCGATG 780
Qy      261 GlyProTyrLyValSerIleGlyIleLysProThrGluGlyLeuMetValAspPheSer 280
Db      781 GGGCGGATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 840
Qy      281 AspIleLeuArg-ProTyrGlyAsnHisThrGlyLeuSerHisLeuPheSerValGluAlaAs 300
Db      841 GATATATCA-ACCATATCGGCAATCATACGGGTAATCTCCGCCATCCGTAAGAGGCTGATA 899
Qy      300 nserHisglGluGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGln 320
Db      900 CAGTCATGAGGGGTTATGATATACAGCATGACAGCATGACAGCATGACAGCATGACAGC 959
Qy      320 o 320
Db      960 T 960

```

```

RESULT 5
AAZ53688
ID      AAZ53688 standard; DNA; 963 BP.
AC      AAZ53688;
XX      15-SEP-2003 (revised)

```

```

DT      21-MAR-2000 (first entry)
XX      Neisseria gonorrhoeae ORF 406 partial DNA sequence SEQ ID NO:1325.
DE      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX      Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW      antigen; diagnosis; immunogenic; infection; meningitis; septicemia;
KW      antibacterial; gene therapy; de.
XX      Neisseria gonorrhoeae.
OS      Neisseria gonorrhoeae.
XX      WO9957280-A2.
XX      11-NOV-1999.
XX      30-APR-1999; 99WO-US009346.
PF      01-MAY-1998; 98US-0083758P.
PR      31-JUL-1998; 98US-0094869P.
PR      02-SEP-1998; 98US-0098994P.
PR      02-SEP-1998; 98US-0099063P.
PR      09-OCT-1998; 98US-0103749P.
PR      09-OCT-1998; 98US-0103794P.
PR      09-OCT-1998; 98US-0103796P.
PR      25-FEB-1999; 99US-0121528P.
XX      (CHIR) CHIRON CORP.
PA      (GENO-) INST GENOMIC RES.
XX      Frazer C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
PI      Petersen J, Pizza M, Rapunoli R, Ratti G, Scalato E, Scarselli M,
PI      Tettelin H, Venter JC.
DR      WPI, 2000-062150/05.
XX      P-PSDB; AAY74926.
XX      Novel Neisserial polypeptides predicted to be useful antigens for
PT      vaccines and diagnosis.
XX      Claim 7, Page 728; 1453pp; English.
XX      AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
XX      represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX      and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
XX      PCR primers used in the exemplification of the present invention. The
XX      polypeptides, the polynucleotides, antibodies and compositions of the
XX      invention can be used as vaccines, as diagnostic reagents, and as
XX      immunogenic compositions. The polypeptides can be used in the manufacture
XX      of medicaments for treating or preventing infection due to Neisserial
XX      bacteria (e.g. meningitis and septicemia), to detect the presence of
XX      Neisseria bacteria, or to raise antibodies. They may also be used to
XX      screen for agonists or antagonists, which may themselves have use as
XX      antibacterial agents. The polynucleotides of the invention may also be
XX      used in gene therapy protocols. (Updated on 15-SEP-2003 to standardise OS
XX      field)
SQ      Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other.

Alignment Scores:
Pred. No.: 1,21e-279 Length: 963
Score: 282.00 Matches: 319
Percent Similarity: 99.38% Conservative: 0
Beet Local Similarity: 99.38% Mismatches: 1
Query Match: 88.12% Indels: 2
DB: 3 Gaps: 0

US-10-617-835-4 (1-320) x AAZ53688 (1-963)
Qy      1 MetArgAlaArgGluLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Db      1 ATGGGGGACGGGCTGCTGATACCTATCTTTTTCAGTTTATTTTATCCGCGCTGGCGG 60
Qy      21 ThrLeuThrGlyIleProSerHisglYglYlyValGPhenAlaValGluGlnGluLeu 40

```

```

Dh 61 ACATGACAGGATATTCATCCATGCGAGGCAAAACGCTTCGCGGTGCAACAAAGACTT 120
Qy 41 ValAlaAlaSerAlaArgAlaAlaValLysPheMetAspLeuGluAlaLeuHisGlyArg 60
Dh 121 GTGGCGGCTTCTGCGAGAGTGGCGTTAAAGACATGATTTACAGGCACTTACACGAGACA 180
Qy 61 LysValAlaLeuLysIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Dh 181 AAAGTTGATTTGATTCATTCGAACATATGGGCGACCAAGGTTACAGGAGTTTACAGGGGGT 240
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100
Dh 241 CGCTACCTCCATTTGATGACCTGATTCGCGCGAATACATTAACAGCCCTGCGCGCACCC 300
Qy 101 AspTyrThrTyrProArgTyrGlyTyrThrAlaGluTyrThrSerGlyGlyLeuThrGly 120
Dh 301 GATTACACTTATCCGCGTTACGAAACACCGCTGAAACACATCAGGCGGTTTACGGGGT 360
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Dh 361 TTAACTCACTTCTTATATCTACACTTAATGCCCTGCACTCTGCGCACCCAAATCAGACGGT 420
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160
Dh 421 AGCGGAAGTAGACACAGCTGCGCTTAAATATGGCGGATGGGGGATTTATCGAAATGAA 480
Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisIleLeuValGlnThrValPhe 180
Dh 481 ACCTTGACGACCAACCCGCGGACACTGCCCTTTCTTTCCCACTTGGTCAAGCCGTATTT 540
Qy 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Dh 541 TTCCTGGCGGCGATAGACGTTGTTTCTCTGCGCAATGCCGATACAGATGTTATTATTAAC 600
Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGlyMetHisIleLeuTyrAsnAlaGluThr 220
Dh 601 ATCAACGATTTGGAAGCATACGCAACGAAACCGAAATGACCTATTAACAATGCCGAAACA 660
Qy 221 LeuLysAlaGlnTyrLeuGlyTyrPheAlaValAspArgThrAsnLysLeuLeu 240
Dh 661 CTGAAAGCCCAACAACTGGAATATTTCCGACATAGACAAACCAATMAAAATTTGCTC 720
Qy 241 IleLysProLysThrAsnAlaPheGlnAlaAlaTyrLysGluAsnTyrAlaLeuTyrMet 260
Dh 721 ATCAAAACCAAAACCAATGCGTTTGAAGCTGCTCTATAAGAAATTAAGCATTTGGGATG 780
Qy 261 GlyProTyrLysValSerLysGlyIleLysProThrGlnGlyLeuMetValAspPheSer 280
Dh 781 GGGCCGATTAAGTAAGCAAGGAATCAACCGACGGAAGGATTTGATGCTGATTTCTCC 840
Qy 281 AspIleArg-ProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAs 300
Dh 841 GATATCCA-ACCATATCGGCAATCATACGGGTAATCTCCGCCCATTCCTATAGAGGCTGATA 899
Qy 300 nSerHisGlnGlyTyrGlyTyrSerAspGluAlaValArgGlnHisArgGlnGlyGlnP 320
Dh 900 CAGTCATGAGGGGATGATACAGCATGAAAGCATGCGACCAACATAGACAGGCGCAACC 959
Qy 320 o 320
Dh 960 T 960

```

RESULT 6  
AAA81297  
ID AAA81297 standard; DNA; 963 BP.

AC AAA81297;  
XX 15-SEP-2003 (revised)  
DT 04-DEC-2000 (first entry)  
XX

N. gonorrhoeae partial DNA sequence g406.seq SEQ ID NO:1036.

```

KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW Meningococcus B; MenB; ds.  
OS Neisseria gonorrhoeae.  
PN W0200022430-A2.  
PD 20-APR-2000.  
PP 08-OCT-1999; 99WO-US023573.  
PR 09-OCT-1998; 98US-0103794P.  
PR 30-APR-1999; 99US-0132068P.  
XX (CHIR ) CHIRON CORP.  
PA  
PI Frazer CM, Hickey E, Peterson J, Tettelein H, Venter JC;  
PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
PI Rappuoli R, Pizza M;  
XX WPI: 2000-318079/27.  
DR P-PSDB; AAB25657.  
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used  
PT in the diagnosis and treatment of N. meningitidis infection and other  
PT Neisseria infections, for example, N.gonorrhoea.  
PS Claim 9; Page 108; 1760pp; English.  
XX  
XX The present invention describes methods of obtaining immunogenic proteins  
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent  
CC specifically claimed Neisseria meningitidis genomic DNA sequences;  
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA  
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and  
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of  
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent  
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all  
CC used in the exemplification of the present invention. The nucleic acid  
CC sequences, protein sequences, and antibodies against them, can be used in  
CC the manufacture of a composition. The composition can be used as a  
CC medicament (or in the manufacture of a medicament) for treating,  
CC preventing or diagnosing infection due to Neisseria bacteria. For  
CC example, some of the identified proteins could be components of vaccines  
CC against Meningococcus B; against all serotypes; and/or against all  
CC pathogenic Neisseriae. Identification of sequences from the bacterium  
CC will also facilitate production of biological probes, particularly  
CC organism-specific probes. Attempts to make efficacious Meningococcus B  
CC vaccines have failed mainly due to antigen tolerance. Multivalent  
CC vaccines have also been tried but none have successfully overcome  
CC antigenic variability. The provision of further, complete sequences may  
CC provide an opportunity to identify secreted or surface exposed proteins  
CC that may be presumed targets for the immune system and which are not  
CC antigenically variable or at least more conserved than other more  
CC variable regions. (Updated on 15-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other;  


```

#### Alignment Scores:

Pred. No.: 1,21e-279 Length: 963  
Score: 282.00 Matches: 319  
Percent Similarity: 99.38% Conservative: 0  
Best Local Similarity: 99.38% Mismatches: 1  
Query Match: 88.12% Indels: 2  
DB: 3 Gaps: 0

US-10-617-835-4 (1-320) x AAA81297 (1-963)

```

Qy 1 MeArGAlaArgLeuLeuIlePheIleLeuPheSerValPheIleLeuSerAlaCysGly 20
Dh 1 ATGCGGACACGGCGTGCATACCTATCTTTTTCACATTTTATTTATCTACGCTCGCGG 60
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyIlyLysArgPheAlaValGluGlnGluLeu 40

```



```

Db      61 ACACGACAGGATTCATCGATGCGGAGCAACCGTTCCGCTCGAACAAGACTT 120
Qy      41 ValAlaAlaSerAlaArgAlaAlaValAlaAspMetAspLeuGlnAlaLeuHisGlyArg 60
Db      121 GTGGCCGCTTGTGGCAGAGCTGCCGTTAAAGACATGATTTACAGGACATTAACCGGACGA 180
Qy      61 LysValAlaLeuTyrTlAlaThreMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80
Db      181 AAAGTTGATGTGTACATTGCAACTATGGCGACCAAGGTTGAGGAGTTTTCAGAGGGGT 240
Qy      81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAspSerProAlaValArgThr 100
Db      241 CGCTACTCTCATTTGATGACACTGATTTCCGCGCGAATCATTAACAGCCCTCCGCTCCGACC 300
Qy      101 AspTyrThrTyrProArgTyrGlyGlnThrAlaGlnThrThrSerGlyGlyLeuThrGly 120
Db      301 GATTACACTTATCCCGTTACGAAACCAACCGCTGAACACATCAAGCGGTTTGAAGGGGT 360
Qy      121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140
Db      361 TTAAACACTTCTTTTATCTACACTTAATGCCCTGCACTCTCCGCAACCCCAATCAACACGT 420
Qy      141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGly 160
Db      421 AGCGGAAGTAGAGAGAGCTGCGCTTAATATTGGCGGAGTGGGAGATTATCGAAATGAA 480
Qy      161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      481 ACCTTGACAGACCAACCCCGCGGACACTGCTTTCTTCCCACTGTGTGACAGCCTATATT 540
Qy      181 PheLeuAlaGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541 TTCCGCGCGGATAGACGTTGTTCTCTCCCAATGCGATACAGATGTGTTATTATAC 600
Qy      201 IleAspValPheGlyThrIleArgAsnArgThrGlnMetHisLeuTyrAsnAlaGlnThr 220
Db      601 ATCGACGATTCGAGACATACGCAACGAAACCGAAATGACCTTATCAATGCCAACA 660
Qy      221 LeuIysAlaGlnThrIleLeuGlnIuTyrPheAlaValAspArgThrAsnIleLeuLeu 240
Db      661 CTGAAGGCCCAACCAACTGGAAATATTTCGAGTAGACAGAACCAATTAATAAATTGCTC 720
Qy      241 IleIysProIysThrAsnAlaPheGlnAlaIleTyrIleGlyLeuTyrAlaLeuTyrMet 260
Db      721 ATCAAAACCAAAACCAATGCGTTGTAACCTGCTTAAAGAAATTAAGCATTTGAGATG 780
Qy      261 GlyProTyrIleValSerIleGlyIleIysProThrGlnGlyLeuMetValAspPheSer 280
Db      781 GGGCGGTATTAAGTAGCAAGAAAGAAATCAACCGAGGAAAGATGATGTCGATTTCTCC 840
Qy      281 AspIleArg-ProTyrGlyAsnHisIleThrGlyAsnSerAlaProSerValGlnAlaAsp 300
Db      841 GATATCCA-ACCATATCGGCAATCATATCGGTAATCCGCCCATCGTAGAGAGGTGATTA 899
Qy      300 nSerHisGlnGlyTyrGlyTyrSerAspGlnAlaIleAlaArgGlnHisArgGlnGlyGln 320
Db      900 CAGTCATAGAGGGATGATGATACAGGATGAAGACATGCGCAACATTAACAAGGCGCAAC 959
Qy      320 o 320
Db      960 T 960

```

RESULT 7  
AAF21582  
ID AAF21582 standard; DNA; 963 BP.  
XX  
XX AAF21582;  
XX AC  
DT 15-SEP-2003 (revised)  
DT 13-MAR-2001 (first entry)  
XX  
N. gonorrhoeae partial DNA sequence g406.seq SEQ ID NO:76.

```

XX      XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;  
KM      KM diagnosis; antigen; detection; infection; gene therapy; antibacterial;  
KW      KW ds.  
XX      XX  
OS      OS Neisseria gonorrhoeae.  
XX      XX  
PN      PN W0200066791-A1.  
XX      XX  
PD      PD 09-NOV-2000.  
XX      XX  
PF      PF 08-MAR-2000; 2000WO-US005928.  
XX      XX  
PR      PR 30-APR-1999; 99US-0132068P.  
PR      PR 08-OCT-1999; 99WO-US023573.  
XX      XX 28-FEB-2000; 2000GB-00004695.  
XX      XX  
PA      PA (CHIR) CHIRON CORP.  
PA      PA (GENO-) INST GENOMIC RES.  
XX      XX  
PI      PI Pizza M, Hickey E, Peterson J, Tetteijn H, Venter JC;  
PI      PI Maignani V, Galeotti C, Mora M, Ratti G, Scarvelli M, Scarlato V;  
PI      PI Rapuoli R, Frazer CM, Grandi G;  
XX      XX  
DR      DR WPI, 2000-647603/62.  
DR      DR P-PEDB; AAB58587.  
XX      XX  
PT      PT Neisseria meningitidis B full length genome sequence and open reading  
PT      PT frames are used to detect, treat and prevent Neisserial infections.  
XX      XX  
PS      PS Example 1; Page 109; 692pp; English.  
XX      XX  
CC      CC The present invention describes the full length genome of Neisseria  
CC      CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613  
CC      CC represent fragments of the NMB genomic sequence, as the sequence was too  
CC      CC long to go in a record on its own it was split into 8 sequences which  
CC      CC overlap each other at the beginning and end of each sequence by 49980 bp  
CC      CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of  
CC      CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of  
CC      CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins  
CC      CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR  
CC      CC primers which are used in the exemplification of the present invention.  
CC      CC The NMB genome and fragments from it have antibacterial activity, and can  
CC      CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins  
CC      CC and/or antibodies which bind to the proteins can be used in compositions  
CC      CC for treating or preventing infection due to Neisserial bacteria or as a  
CC      CC diagnostic reagent for detecting the presence of Neisserial bacteria or  
CC      CC of antibodies raised to Neisserial bacteria. Computers, computer memory,  
CC      CC computer storage medium or computer databases can be used in a search to  
CC      CC identify open reading frames (ORFs) or coding sequences within the NMB  
CC      CC genome. The DNA sequences provide further opportunities to find antigenic  
CC      CC or immunogenic proteins which are more effective in vaccines than the  
CC      CC outer membrane proteins currently used. (Updated on 15-SEP-2003 to  
XX      XX standardise OS field)  
SQ      SQ Sequence 963 BP; 276 A; 241 C; 226 G; 220 T; 0 U; 0 Other:  
XX      XX  
Alignment Scores:  
Pred. No.: 1,21e-279 Length: 963  
Score: 282.00 Matches: 319  
Percent Similarity: 99.38% Conservative: 0  
Best Local Similarity: 99.38% Mismatches: 1  
Query Match: 88.12% Indels: 2  
DB: 3 Gaps: 0  
US-10-617-835-4 (1-320) x AAF21582 (1-963)  
Qy      1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGly 20  
Db      1 ATGCGGCGACGGCTGCTGATACCTATTCCTTTTTCAGTTTATTTATTCGCGCGG 60  
Qy      21 ThrLeuThrGlyIleProSerHisGlyGlyIleIysAspArgPheAlaValGlnGlnIleu 40

```

Db 61 ACACGACAGGATTCATCGCATGCGAGGACAAACGCTTCGCGGTCAACAAAGACTT 120  
Qy 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
Db 121 GTGGCCGCTTCCTGCGAGAGCTGCGTTAAAGACATGATTTACAGGCACTTACACGACGA 180  
Qy 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
Db 181 AAAGTGCATTTGATTCATTCGCAACTATGGCGACCAAGATTACAGGCACTTACAGGGGT 240  
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
Db 241 CGCTACTCCATTCATGATGCACTGATTCGCGGCAATACATAAACAGCCCTGCGCGCACCC 300  
Qy 101 AspTyrThrTyrProArgTyrGlyThrThrAlaGlyuThrThrSerGlyGlyLeuThrGly 120  
Db 301 GATTACACCTTATCCGCGTTACGAACACCGCTGAACACACATCCAGGCGGTTTGCGGGT 360  
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
Db 361 TTAAACACTTCTTATATCTACACTTATATGCCCTGCACTCGCGCACCCAAATCAGACGCT 420  
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160  
Db 421 AGCGAAGTAGAGAGAGCTGCGGCTTAAATATGGCGGAGTGGGGGATTCGAAATGAA 480  
Qy 161 ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisIleValGlnThrValPhe 180  
Db 481 ACCTTGACGACCAACCCGCGGACACTGCTTCTTCCACTGGTGCACAGCGTATTT 540  
Qy 181 PheLeuArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200  
Db 541 TTCTCGGCGGCGCATAGACGTTGTTCTCTCGCAATCGCATACGATGCTTATATTAC 600  
Qy 201 IleAspValPheGlyThrIleArgAsnArgThrGlyMetHisIleuTyrAsnAlaGlyThr 220  
Db 601 ATCCACGATTCGGAACGATACGCAACAGAACCCGAAATGACCTATACAAATGCCGAAACA 660  
Qy 221 LeuTyrAlaGlnThrTyrLeuGlyuTyrPheAlaValAspArgThrAsnLysLeuLeu 240  
Db 661 CTGAAAGCCCAACAAACCTGAAATATTTCCGACGTAGACAGAACCAATTAATAATTGCTC 720  
Qy 241 IleLysProLysThrAsnAlaPheGlyAlaAlaIleTyrLysGlyAsnTyrAlaLeuTrpMet 260  
Db 721 ATCAAAACCAAAACCAATGCGTTGAACTGCTCATTAAGAAATTTACGCACTTGCGATG 780  
Qy 261 GlyProTyrLysValSerLysGlyIleLysProThrGlyGlyLeuMetValAspPheSer 280  
Db 781 GGGCGGATTAAGTAAGCAAGAAATCAACCGAAGGATTCGATGCTGATTTCTCC 840  
Qy 281 AspIleArg-ProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGlyAlaAspAs 300  
Db 841 GATATCCA-ACCATAACGGCAATCATACGGGTAACTCCGCCCAATCCGTAAGGCTGATTA 899  
Qy 300 nSerHisGlyGlyTyrGlyTyrSerAspGlyAlaValArgGlnHisArgGlnGlyInPr 320  
Db 900 CAGTCATGAGGGGATGATGATACAGGATGAGAGCAAGTGGCAACATAGACAAAGGCAACC 959  
Qy 320 C 320  
Db 960 T 960

RESULT 8  
ABZ40799  
ID ABZ40799 standard; DNA; 789 BP.  
XX ABZ40799;  
XX  
XX 07-MAR-2003 (first entry)  
DT  
XX  
DE N. gonorrhoeae nucleotide sequence SRQ ID 6187.  
XX  
XX Antibacterial; infection; vaccine; gene therapy; gene; ds.

XX OS Neisseria gonorrhoeae.  
XX EN W0200279243-A2.  
XX PD 10-OCT-2002.  
XX PF 12-FEB-2002; 2002WO-1B002069.  
XX PR 12-FEB-2001; 2001GB-00003424.  
XX PA (CHIR-) CHIRON SPA.  
XX PI Fontana MR, Pizsa M, Masiagnani V, Monaci E;  
XX WPI: 2003-058415/05.  
XX DR P-PSDB; ABP79829.  
XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
XX PT medicament for treating or preventing N. gonorrhoeae infection.  
XX PS Disclosure; Page 639; 815pp; English.  
XX CC The present invention relates to proteins from Neisseria gonorrhoeae.  
XX CC Also disclosed are the nucleic acid molecules encoding the proteins and  
XX CC antibodies that specifically bind to the proteins. The composition  
XX CC comprising the protein, nucleic acid or antibody is useful for the  
XX CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
XX CC infection, this may be in the form of a vaccine or gene therapy.  
XX CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid  
XX CC molecules of the invention  
XX SQ Sequence 789 BP; 217 A; 203 C; 181 G; 188 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.91e-254 Length: 789  
Score: 257.00 Matches: 257  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 80.31% Indels: 0  
DB: Gaps: 0  
US-10-617-835-4 (1-320) x ABZ40799 (1-789)  
Qy 1 MetArgAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCyGly 20  
Db 1 ATGGGGGACAGCGCTGCTGATACCATTTCTTTTCACTTTTATTATCGCGCTGGGG 60  
Qy 21 ThrLeuThrGlyIleProSerHisGlyGlyGlyLysArgPheAlaValGlnGlnLeu 40  
Db 61 ACACGACAGGATTCATTCATGCGCATGGCGGACAAACGCTTCGCGGTCAACAAAGACTT 120  
Qy 41 ValAlaAlaSerAlaArgAlaAlaValLysAspMetAspLeuGlnAlaLeuHisGlyArg 60  
Db 121 GTGGCCGCTTCCTGCGAGAGCTGCGTTAAAGACATGATTTACAGGCACTTACAGGACGA 180  
Qy 61 LysValAlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyGly 80  
Db 181 AAAGTGCATTTGATTCATTCGCAACTATGGCGACCAAGATTACAGGCACTTACAGGGGT 240  
Qy 81 ArgTyrSerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThr 100  
Db 241 CGCTACTCCATTCATGATGCACTGATTCGCGGCAATACATAAACAGCCCTGCGCGCACCC 300  
Qy 101 AspTyrThrTyrProArgTyrGlyuThrThrAlaGlyuThrThrSerGlyGlyLeuThrGly 120  
Db 301 GATTACACCTTATCCGCGTTACGAACACCGCTGAACACACATACAGGCGGTTGACGGGT 360  
Qy 121 LeuThrThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGly 140  
Db 361 TTAAACACTTCTTATATCTACACTTATATGCCCTGCACTCGCGCACCCAAATCAGACGCT 420  
Qy 141 SerGlySerArgSerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGlu 160

```

Db      421  AGCGAAGTAGAGCGAGTCTGGGCTTAATATTTGGCGGATGGGGATTAATGAAATGA 480
Qy      161  ThrLeuThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhe 180
Db      481  ACCGTGACGACCAACCCGGCGACACTGCTTCTTCCCACTTGSTGCAGACCGTATTT 540
Qy      181  PheLeuArgGlyTlleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsn 200
Db      541  TTCCTGGCGGATAGACGTTGTTCTCTGCGCAATGCGATACAGATGTGTTATTAAAC 600
Qy      201  IleAspValPheGlyThrIleArgAsnArgThrGluMetHisLeuThrAsnAlaGluThr 220
Db      601  ATCGACGATTTTCGGAACGATACGCAACAGAACCGAAATGCACTATACATGCCGAAACA 660
Qy      221  LeuValAlaGlnThrIleLeuGluThrPheAlaValAspArgThrAsnValLeuLeu 240
Db      661  CTGAAGGCCCAACAAACCTGGAATATTTCCGAGTAGACAGAACCAATAAAAATTGCTC 720
Qy      241  IleValProValThrAsnAlaPheGluAlaIleTyrIleGluAsnTyrAla 257
Db      721  ATCAAAACCCAAACCAATGCTTTGAGAGCTGCTATTAAGAAATTATGACC 771

```

## RESULT 9

AA212006 standard, DNA; 963 BP.

AA212006;

08-OCT-1999 (first entry)

Neisseria meningitidis strain B complete ORF15 sequence.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine; treatment; Neisseria infection; meningitis; septicemia; gonorrhea; ss.

Neisseria meningitidis.

MO9924578-A2.

20-MAY-1999.

09-OCT-1998; 98WO-1B00165.

06-NOV-1997; 97GB-00023516.

14-NOV-1997; 97GB-00024190.

27-NOV-1997; 97GB-00024386.

10-DEC-1997; 97GB-00025158.

14-JAN-1998; 98GB-00000759.

01-SEP-1998; 98GB-00019016.

(CHIR-) CHIRON SPA.

Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;

WPI; 1999-327407/27.

P-PSDB; AAY38538.

Proteins from Neisseria meningitidis and N. gonorrhoeae useful for diagnosis, treatment and prevention of infection.

Claim 9, Page 98; 524pp; English.

Nucleotide sequences AA211972-212358 represent open reading frames (ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their fragments, their nucleic acids and antibodies are used for diagnosis, prevention (as vaccines) or treatment of Neisseria infections, such as meningitis, septicemia and gonorrhea. Both organisms are closely related. Fragments of the nucleic acids are useful as hybridisation probes and antisense reagents

Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1,46e-202 Length: 963  
Score: 207.00 Matches: 307  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 1  
Query Match: 64.69% Indels: 2  
DB: 2 Gaps: 0

US-10-617-835-4 (1-320) x AA212006 (1-963)

```

Qy      3  AlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGlyThrLeu 22
Db      7  GCACGGCTGCTGATACCTTATTTCTTTTTCAGTTTATTATTATTCGCCCTGGGACACAG 66
Qy      23  ThrGlyIleProSerHisGlyGlyGlyValArgPheAlaValGluGlnIleLeuValAla 42
Db      67  ACAGGATTCATCCGATGCGGAGGTAAACGCTTGGCGTGAACAAAGAACTGTGGCC 126
Qy      43  AlaSerAlaArgAlaAlaValAlaAspMetAspLeuGlnAlaLeuHisGlyArgVal 62
Db      127  GCTTCTGCCAGAGCTGCGCTTAAGACATGATTTTACAGGCATTTACACGACGAAAGTT 186
Qy      63  AlaLeuTyrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlyArgTyr 82
Db      187  GCATGTACATTTGCCATCTATGCGGCGACCAAGGTTTCAGCATTTTACAGGGGCTGCTAC 246
Qy      83  SerIleAspAlaLeuIleArgGlyGlyTyrIleAsnSerProAlaValArgThrAspTyr 102
Db      247  TCATTTGATGACACTGATGTTGGCGAATACATTAACAGCCCTGCGCTACCGATAC 306
Qy      103  ThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThr 122
Db      307  ACCCTATCCACGTTAGGAAACCAACCGCTGAACATCGAGGCTTTCAGAGTTTAAACC 366
Qy      123  ThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142
Db      367  ACTTCTTTATCTACCTTATATGCCCCCTGCACTCTTCGCAACCCATCAGACGGTAGCGGA 426
Qy      143  SerArg-SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGluThrLe 162
Db      427  AGTAA-AAGCAGCTGGGGCTTAATATTTGCGGATGGGGGATTTACGAATGAACCTT 485
Qy      162  UthrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePhe 182
Db      486  GACGACTAACCCGCGCGACACTGCTTCTTCCCACTTGTAACAGACCGTATTTTCTCT 545
Qy      182  ValArgIleIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsnIleAs 202
Db      546  GCGGCGCATAGACGTTGTTTCTCTGCGCAATGCCATACAGATGTGTTTATTACATGCA 605
Qy      202  PValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThrLeu 222
Db      606  CGTATTCGGAACGATACGCAACAGAACCGAAATGCACTATCAATGCGGAACACTGAA 665
Qy      222  ValGlnThrIleValLeuGluThrPheAlaValAspArgThrAsnValLeuLeuIle 242
Db      666  AGCCCAACCAAACTCGAATATTTGCGAGTACAGAACCAATTAATAAATTTCTCATCA 725
Qy      242  AspProValThrAsnAlaPheGluAlaAlaTyrIleGluAsnTyrAlaLeuTyrMetGly 262
Db      726  ACCCAAAACCAATGCGTTTGAAGCTGCTTATTAAGAAATTTACGATTTGTGATGGGGCC 785
Qy      262  OTyrIleValSerIleValLeuProThrGluGlyLeuMetValAspPheSerAsp 282
Db      786  GTATTAAGTAAGCAAAAGGATTAACCGACGGAAGATTAAGTGCGATTTCTCGCATAT 845
Qy      282  ValArgProTyrGlyValAsnHisThrGlyAsnSerIleProSerValGluAlaAspAsnSer 302
Db      846  CCGACCATACGGCAATCATACGGGTAACTCCGCCCATTCGTAAAGGCTGATTAACAGTCA 905
Qy      302  GGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 310

```



```

XX DE Neisseria meningitidis ORF 406 partial DNA sequence SEQ ID NO:3109.
XX XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX KW antibacterial; gene therapy; ds.
XX XX
XX OS Neisseria meningitidis.
XX XX
XX PN WO957280-A2.
XX PD 11-NOV-1999.
XX PF 30-APR-1999; 99WO-US009346.
XX XX
XX PR 01-MAY-1998; 98US-0083758P.
XX PR 31-JUL-1998; 98US-0094869P.
XX PR 02-SEP-1998; 98US-0098994P.
XX PR 02-SEP-1998; 98US-0099062P.
XX PR 09-OCT-1998; 98US-0103749P.
XX PR 09-OCT-1998; 98US-0103794P.
XX PR 09-OCT-1998; 98US-0103796P.
XX PR 25-FEB-1999; 99US-0121528P.
XX XX
XX PA (CHIR ) CHIRON CORP.
XX PA (GENO-) INST GENOMIC RES.
XX PI Frazer C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX PI Petersen J, Pizze M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
XX PI Tetrelin H, Venter JC;
XX XX
XX DR MPI: 2000-062150/05.
XX DR P-PSDB: AAY75811.
XX PT Novel Neisserial polypeptides predicted to be useful antigens for
XX PT vaccines and diagnostics.
XX XX
XX PS Example 1; Page 108-109; 1453pp; English.
XX XX
XX CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
XX CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
XX CC PCR primers used in the exemplification of the present invention. The
XX CC polypeptides, the polynucleotides, antibodies and compositions of the
XX CC invention can be used as vaccines, as diagnostic reagents, and as
XX CC immunogenic compositions. The polypeptides can be used in the manufacture
XX CC of medicaments for treating or preventing infection due to Neisserial
XX CC bacteria (e.g. meningitis and septicemia), to detect the presence of
XX CC Neisseria bacteria, or to raise antibodies. They may also be used to
XX CC screen for agonists or antagonists, which may themselves have use as
XX CC antibacterial agents. The polynucleotides of the invention may also be
XX CC used in gene therapy protocols
XX XX
XX SQ Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;
XX XX
XX Alignment Scores:
XX Pred. No.: 1,466-202 Length: 963
XX Score: 207.00 Matches: 307
XX Percent Similarity: 99.35% Conservativity: 0
XX Best Local Similarity: 99.35% Mismatches: 1
XX Query Match: 64.69% Indels: 2
XX DB: 3 Gaps: 0
XX
XX US-10-617-835-4 (1-320) x AA254613 (1-963)
XX
QY 3 AAlaAGLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCySGlyThrLeu 22
DB 7 GCAAGCGCTGCTATACCTATTTCTTTTTCAGTTTATTTATTTATTCGCCCGCGGACACTG 66
QY 23 ThrGlyIleProSerHisGlyGlyGlyIleValArgPheAlaValGluGlnGluLeuValAla 42.
DB 67 ACAGGTATTCATCCATGCGAGGTAAACGCTTTGCGGTGGAACAGAACTGTGGCC 126

```

```

QY 43 AAlaSerAlaArgAlaAlaValIleValPheMetAspLeuGlnAlaLeuHisGlyValArgIleVal 62
DB 127 GCTTCTGCGAGAGCTGCGGTAAAGACATGATTAACAGCATTTACACGAGCAAAAGTT 186
QY 63 AAlaLeuValIleAlaThrMetGlyValArgGlnGlySerGlySerLeuThrGlyValArgTyr 82
DB 187 GCATTGTACATTGCCACATATGGGCGACCAAGGTTTACAGGAGTTTACACAGGGGTCCTAC 246
QY 83 SerIleAspAlaLeuIleArgGlyGlyValIleValSerProAlaValArgThrAspTyr 102
DB 247 TCCATTGATGACACTGATTCGTGGCGAATACATTAACAGCCCTGCGCTCGATCCCATTTAC 306
QY 103 ThrTyrProArgTyrGlyThrThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThr 122
DB 307 ACCATACCAAGTTACGAAACCAACCGCTGAAACATACAGCGGTTTACAGGTTTAAACC 366
QY 123 ThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142
DB 367 ACTTCTTATCTACACTTAATCCCTGTGACTCTCTCGACCCATACAGCGGTAGCGGA 426
QY 143 SerArg-SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrArgAsnGluThrIle 162
DB 427 AGTAA-AAAGCAAGTCTGGGCTTAATATTTGGCGGATGGGGGATTTACGAATGAACCTT 485
QY 163 uThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePhe 182
DB 486 GACGACTAACCCGCGAGACTGCTTTCTTCCACTTGTGACAGACCGTATTTTCT 545
QY 182 uArgGlyIleAspValIleSerProAlaAsnAlaAspThrAspValPheIleAsnIleAs 202
DB 546 GCGCGGCAATAGACGTTGTTTCTCTGCAATCCGATACAGATGTTTATTAACATGCA 605
QY 202 pValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThrLeu 222
DB 606 CGTATTCCGAACGATACCAACGAAACGAAATGCACTATCAATGCGGAAACACTGAA 665
QY 222 sAlaGlnThrIleValLeuGluTyrPheAlaValAspArgThrAsnIleValLeuIle 242
DB 666 AGCCCAACCAAACTGGAATATTTGCAGTACACGAACCAATTAATAATTTCTCATCA 725
QY 242 sProIleThrAsnAlaPheGluAlaAlaTyrIleValGlnAsnTyrAlaLeuTyrMetGly 262
DB 726 ACCAATAAACCAATGCTTTGAAGCTGCTATTAAGAAATTAACGATTTGTGATGGGCGC 785
QY 262 cTyrIleValSerIleValIleValProThrGluGlyLeuMetValAspPheSerAsp 282
DB 786 GTATTAAGATTAAGCAAGAAATTAACCGACGAAAGATTAATGTCGATTTCTCCGATAT 845
QY 282 eArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluIleAspAsnSer 302
DB 846 CCGACCATACCGCAATCATACGGGTAACTCCGCCCATTCGTAAGGCTGATTAACATGCA 905
QY 302 sGluGlyTyrGlyTyrSerAspGlu 310
DB 906 TGAGCGGATGATGATACAGCATGAA 930
XX
XX RESULT 12
XX AAA81296
XX ID AAA81296 standard; DNA; 963 BP.
XX XX
XX AC AAA81296;
XX XX
XX DT 04-DEC-2000 (first entry)
XX XX
XX DE N. meningitidis partial DNA sequence m406.seq SEQ ID NO:1034.
XX XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
XX KW Meningococcus B; Memb; ds.
XX XX
XX OS Neisseria meningitidis.
XX XX
XX PN WO200022430-A2.

```

```

XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US023573.
XX
XX 09-OCT-1998; 98US-0103794P.
XX 30-APR-1999; 99US-0132068P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tettein H, Venter JC,
XX Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,
XX Rappuoli R, Pizza M;
XX WPI; 2000-318079/27.
XX P-PSDB; AAB25656.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
XX in the diagnosis and treatment of N. meningitidis infection and other
XX Neisserial infections, for example, N.gonorrhoea.
XX
XX Claim 9; Page 107; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
XX from Neisseria genomic sequences. AAA81453 to AAA82414 represent
XX specifically claimed Neisseria meningitidis genomic DNA sequences;
XX AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
XX sequences and their corresponding proteins; AAA81254 to AAA81259 and
XX AAA81304 to AAA81321 represent PCR primers used in the isolation of
XX Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
XX Neisseria meningitidis MemB polynucleotide ORF sequences, which are all
XX used in the exemplification of the present invention. The nucleic acid
XX sequences, protein sequences, and antibodies against them, can be used in
XX the manufacture of a composition. The composition can be used as a
XX medicament (or in the manufacture of a medicament) for treating
XX preventing or diagnosing infection due to Neisserial bacteria. For
XX example, some of the identified proteins could be components of vaccines
XX against Meningococcus B; against all serotypes; and/or against all
XX pathogenic Neisseriae. Identification of sequences from the bacterium
XX will also facilitate production of biological probes, particularly
XX organism-specific probes. Attempts to make efficacious Meningococcus B
XX vaccines have failed mainly due to antigen tolerance. Multivalent
XX vaccines have also been tried but none have successfully overcome
XX antigenic variability. The provision of further, complete sequences may
XX provide an opportunity to identify secreted or surface exposed proteins
XX that may be presumed targets for the immune system and which are not
XX antigenically variable or at least more conserved than other more
XX variable regions
XX
XX SQ Sequence 963 BP; 284 A; 234 C; 217 G; 228 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,46e-202 Length: 963
XX Score: 207.00 Matches: 307
XX Percent Similarity: 99.35% Conservative: 0
XX Best Local Similarity: 99.35% Mismatches: 1
XX Query Match: 64.69% Indels: 2
XX DB: 3 Gaps: 0
XX
XX US-10-617-835-4 (1-320) x AAA81296 (1-963)
XX
XX QY 3 AAlaArgLeuLeuIleProIleLeuPheSerValPheIleLeuSerAlaCysGlyThrIleu 22
XX Db 7 GCACGGCGCTGATACCTATTCTTTTTCAGTTTATTATTATCCGCTGGGGACACTG 66
XX
XX QY 23 ThrGlyIleProSerHisGlyGlyGlyIlyAsnArgPheAlaValIGluGlnGluLeuValAla 42
XX Db 67 ACAGGTATTTCATCGCATGGCGAGGTAAACGCTTGGGCGCAACAAGAACTTGTCGCC 126
XX
XX QY 43 AlaserAlaArgAlaAlaValIlyAsnPheSerPheLeuGlnAlaLeuHisGlyAlaArgIlyVal 62
XX Db 127 GCTTCTGCCAGAGCTGCCGTTAAAGACATGATTTTACAGGCAATTACACGGACGAAAGTT 186

```

```

QY 63 AAlaLeuTrIleAlaThrMetGlyAspGlnGlySerGlySerLeuThrGlnGlyArgTyr 82
Db 187 GCATTGTATACATTGGCAGCATATGGCGACCAAGTTTCAGGCAAGTTTGACAGGGGGTCCGCTAC 246
QY 83 SerIleAspAlaLeuIleArgGlyGluTyrIleAsnSerProAlaValArgThrAspTyr 102
Db 247 TCCATTGATGACATGATTCGTGGCGAATACATAAACGCCCTGCGCCGATCGAATAC 306
QY 103 ThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThr 122
Db 307 ACCTATCCAGCTTTCGAAACCAACCGCTGAACACATCAGCGGTTTGACAGGTTTAAAC 366
QY 123 ThrSerLeuSerThrIleAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142
Db 367 ACTCTTTATATCATCTTAATAGCCCTGCACTCTCTCGACCCATCAGACGGTAGCGGA 426
QY 143 SerArg-SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyrTrpAsnGluThrIle 162
Db 427 AGTAA-AAGCAGTCTGGGCTTAAATATTTGGCGGATGGGATTAATGAAATGAACCTT 485
QY 162 uThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePheLe 182
Db 486 GACGACTAACCCGCGCGACACTGCTTCTTTCCCACTTGGTACAGACCGATTTTTCCT 545
QY 182 uArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsnIleAs 202
Db 546 GCGGCGCATAGACGTTGTTCTCTCGCAATGCGCATACAGATGTGTTATTACATCGA 605
QY 202 pValPheGlyThrIleArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThrIleu 222
Db 606 CGTATTTGGAAACGATACCAACCAACCAATGCACTTATACATGCCGAACACTGAA 665
QY 222 sAlaGlnThrTrpLeuGluTyrPheAlaValAspArgThrAsnIlySerLeuLeuIle 242
Db 666 AGCCCAACAAACACTGAAATATTTCCGATAGACAGAACCAATAAATTTGCTCATCAA 725
QY 242 sProIyThrAsnAlaPheGluAlaAlaTyrIlyGlyAsnTyrAlaLeuTrpMetGlyPr 262
Db 726 ACCAAACCAATGCGTTTGAAGCTGCTTAAAGAAATTTACCATTTGTGGATGGGGCC 785
QY 262 oTyrIlyValSerTrpGlyIleIlyProThrGlnGlyLeuMetValAspPheSerAspI 282
Db 786 GTATTAAGTAAAGCAAGAAATTAACCGACGAAAGATTAATGATTCGATTTCTCCGATAT 845
QY 282 eArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerHi 302
Db 846 CCGACATACGGCATATATACGGGTACTCCGCCCATCGGTAGAGCTGATTAACATGCA 905
QY 302 sGluGlyTyrGlyTyrSerAspGlu 310
Db 906 TGAAGGGTATGATACAGCGATGAA 930
XX
XX RESULT 13
XX AAF21581
XX ID AAF21581 standard; DNA; 963 BP.
XX AC AAF21581;
XX AC
XX DT 13-MAR-2001 (first entry)
XX
XX N. meningitidis partial DNA sequence m406.seq SEQ ID NO:74.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX diagnosis; antigen; detection; infection; gene therapy; antibacterial;
XX ds.
XX
XX Neisseria meningitidis.
XX
XX OS
XX XX
XX PN WC200066791-A1.
XX PD 09-NOV-2000.
XX PF 08-MAR-2000; 2000WO-US005928.

```



PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V,  
PI Rappuoli R, Pizza M,  
XX WPI, 2000-318079/27.  
DR  
XX  
XX Isolated nucleotide sequences of *Neisseria meningitidis* which can be used  
PT in the diagnosis and treatment of *N. meningitidis* infection and other  
PT *Neisseria* infections, for example, *N. gonorrhoea*.  
XX  
XX Claim 7, Page 471-498; 1760pp; English.  
XX  
XX The present invention describes methods of obtaining immunogenic proteins  
CC from *Neisseria* genomic sequences. AAA81453 to AAA82414 represent  
CC specifically claimed *Neisseria meningitidis* genomic DNA sequences;  
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent *Neisseria* DNA  
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and  
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of  
CC *Neisseria meningitidis* DNA sequences; and AAA81322 to AAA81452 represent  
CC *Neisseria meningitidis* MemB polynucleotide ORF sequences, which are all  
CC used in the exemplification of the present invention. The nucleic acid  
CC sequences, protein sequences, and antibodies against them, can be used in  
CC the manufacture of a composition. The composition can be used as a  
CC medicament (or in the manufacture of a medicament) for treating,  
CC preventing or diagnosing infection due to *Neisseria* bacteria. For  
CC example, some of the identified proteins could be components of vaccines  
CC against *Meningococcus B*; against all serotypes; and/or against all  
CC pathogenic *Neisseriae*. Identification of sequences from the bacterium  
CC will also facilitate production of biological probes, particularly  
CC organism-specific probes. Attempts to make efficacious *Meningococcus B*  
CC vaccines have failed mainly due to antigen tolerance. Multivalent  
CC vaccines have also been tried but none have successfully overcome  
CC antigenic variability. The provision of further, complete sequences may  
CC provide an opportunity to identify secreted or surface exposed proteins  
CC that may be pressumed targets for the immune system and which are not  
CC antigenically variable or at least more conserved than other more  
CC variable regions  
XX  
XX Sequence 92934 BP; 24051 A; 24249 C; 21902 G; 22730 T; 0 U; 2 Other;  
SQ

Alignment Scores:  
Pred. No.: 3.09e-187 Length: 92934  
Score: 194.00 Matches: 307  
Percent Similarity: 99.03% Conservative: 0  
Best Local Similarity: 99.03% Mismatches: 1  
Query Match: 60.62% Indels: 3  
DB: Gaps: 0

US-10-617-835-4 (1-320) x AAA81473 (1-92934)

QY 3 AlaargleuLeuileProileLeuPheSerValPheileuSerAlaCYeGlyThrleu 22  
Db 18031 GCAGGGCGCGTGAATCCATCTTTTTCAGTTTATTTATTCGCGCTGGGGACACTG 18090

QY 23 ThrGlyIleProSerHisGlyGlyGlyLysArgPheAlaValGluGlnGluLeuValAla 42  
Db 18091 ACAGGATTTCCATGCGATGCGGAGGTAAAGCTTTGGGTGCGAACAAGAACTTGTGGCC 18150

QY 43 AlaSerIlaArgAlaIaValIysAspMetAspIleuGlnAlaIleuHisGlyArgIysVal 62  
Db 18151 GCTTCTGCAGAGCGCGCTGAAGAACATGGAATTTACGGGATTAACGGGCGAAAGTT 18210

QY 63 AlaLeuTyrlIleAlaThrMetGlyAspGlnIysSerGlySerLeuThr-GlyGlyArgTy 82  
Db 18211 GCATTGTACATTGCCACATATGGCGGACCAAGTTTCAGGCACTTTGACAGGGGGGTCCGTA 18270

QY 82 rSerileAspAlaLeuileArgGlyGlyGlyTyrlIleAsnSerProAlaValArgThrAspTy 102  
Db 18271 CTCATGTGATGACATGATCTGTGGGAATACATAAAGCCCTGCCGTCGATACGATTA 18330

QY 102 rThrTyProArgGlyTyrlThrThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuTh 122  
Db 18331 CACCTATCCAGTTTACGAAACACCGCTGAAACAACATCAAGCGGCTTTGACAGGTTTAAC 18390

QY 122 rThrSerLeuSerThrIleuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerG 142  
Db 18391 CACTTCTTTATCTACATTAATGCGCTGCACTCTCTCGACCAACCAATCAAGCGGTAGCGG 18450

QY 142 ySerArg-SerSerLeuGlyLeuAsnIleGlyGlyMetGlyAspTyArgAsnGluThrL 162  
Db 18451 AAGTAA-AAGAGCTCTGGGCTTAATATTTGGCGGAGGGGATTAATCGAAATGAACCT 18509

QY 162 eUThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGluThrValPhePheL 182  
Db 18510 TGACGACTAACCCCGCGGACACTGCTCTTTCTTCCACTTGATGACAGCGTATTTTCC 18569

QY 182 euArgGlyIleAspValValSerProAlaAsnAlaAspThrAspValPheIleAsnIleA 202  
Db 18570 TGCGCGCATAGACGTTGTTCTCTCTCGCATGCGGATCGATGATGTTATTAATCAATCG 18629

QY 202 spValIlePheGlyThrIleArgAsnArgThrGluMetHisLeuTyrlAsnAlaGluThrleuL 222  
Db 18630 ACGTATTCGAAACATGCGAAGCAAGAACCGAAATGCACCTATACAAATGCGAAACACTGA 18689

QY 222 ySalGlnThrIlyLeuGlyTyrlPheAlaValAspArgThrAsnIlyLeuLeuIleL 242  
Db 18690 AAGCCAAACAAACCTGGAATATTTCCGACTAGACAGAACCAATMAAAATGCTCATCA 18749

QY 242 ySProlYThrAsnAlaPheGluAlaAlaTyrlYsgIuAsnTyrlAlaLeuITrMetGlyP 262  
Db 18750 AACCMAAACCAATGCGTTTGAAGCTGCTATTAAGAAATTAAGCATTTGTGATGGGGCC 18809

QY 262 rOTyrlYsValSerIysGlyIlyleYsProThrGluGlyLeuMetValAspPheSerAspI 282  
Db 18810 CGTTTAAAGTAAGCAAGAAAGAAATTAACCGACGGAAGATTAAAGTGTGATTTCTCCATA 18869

QY 282 leArgProTyrlYsAsnHisThrGlyAsnSerIlaProSerValGluAlaAspAsnSerH 302  
Db 18870 TCCGACCATACCGGCAATCAATACGGGTACTCCGCCCATCCGTAAGAGCTGATTAACGTC 18929

QY 302 lsgIuGlyTyrlYsGlyTyrlSerAspGlu 310  
Db 18930 ATGAGGGGTATGATACAGCGATGAA 18955

RESULT 15  
AAA81489\_7  
Continuation (8 of 9) of AAA81489 from base 700001 (N. meningitidis partial DNA sequence  
WP Sequence split into 9 fragments LOCUS AAA81489 Accession Aaa81489  
WP Fragment Name Begin End  
WP AAA81489\_0 1 110000  
WP AAA81489\_1 100001 210000  
WP AAA81489\_2 200001 310000  
WP AAA81489\_3 300001 410000  
WP AAA81489\_4 400001 510000  
WP AAA81489\_5 500001 610000  
WP AAA81489\_6 600001 710000  
WP AAA81489\_7 700001 810000  
WP AAA81489\_8 800001 837096

Alignment Scores:  
Pred. No.: 3.65e-187 Length: 110000  
Score: 194.00 Matches: 307  
Percent Similarity: 99.03% Conservative: 0  
Best Local Similarity: 99.03% Mismatches: 1  
Query Match: 60.62% Indels: 3  
DB: Gaps: 0

US-10-617-835-4 (1-320) x AAA81489\_7 (1-110000)

QY 3 AlaargleuLeuileProileLeuPheSerValPheileuSerAlaCYeGlyThrleu 22  
Db 86469 GCAAGGCTGCTGATCACTATCTTTTTCAGTTTATTTATTCGCGCTCGGAGACACTG 86528

QY 23 ThrGlyIleProSerHisGlyGlyGlyLysArgPheAlaValGluGlnGluLeuValAla 42  
Db 86529 ACAGGATTTCCATGCGATGCGGAGGTAAAGCTTTGGGTGCGAACAAGAACTTGTGGCC 86588



```
QY      43  A1AserAlaArgAlaIaValIysAspMetAspLeuGlnAlaLeuHisGlyValArgLysVal 62
DB      86589 GCTTCTGCCAAGCTGCGGATTAAGACATGGATTTCACAGCAATTACACGACGAAAGTT 86648
QY      63  A1aLeuTyrI1eAlaThrMetGlyAspGlnGlySerGlySerLeuThr-GlyGlyArgTyr 82
DB      86649 GCATTGTACATTGGCACAATGGGCGCACCAAGGTTCAAGGCAAGTTTGACAGGGGGGTGCTA 86708
QY      82  rSerI1eAspAlaLeuI1eArgGlyGluTyrI1eAsnSerProAlaValArgThrAspTyr 102
DB      86709 CTCCATTGATGCACGATTCTGGGGAATACATAAACGCCCTGCCCTCCGTACCGATTA 86768
QY      102 rThrTyrProArgTyrGluThrThrAlaGluThrThrSerGlyGlyLeuThrGlyLeuThr 122
DB      86769 CACCTATCCAGCTTACGAAACCAACCGCTGAACACATCAGCGGTTTGACAGGTTTAAAC 86828
QY      122 rThrSerLeuSerThrLeuAsnAlaProAlaLeuSerArgThrGlnSerAspGlySerGly 142
DB      86829 CACTTCTTTATCTACACTTAATGCCCTGCACTCTCTCGACCCCAATCAGACGGTAGCGG 86888
QY      142 ySerArg-SerSerLeuGlyLeuAsnI1eGlyGlyMetGlyAspTyrArgAsnGluThrL 162
DB      86889 AAGTAA-AAGCAGTCTGGGCTTAAATATTGGCGGATGGGGATTAATCGAAATGAACCT 86947
QY      162 eutThrThrAsnProArgAspThrAlaPheLeuSerHisLeuValGlnThrValPhePheL 182
DB      86948 TGACGACTTAACCCGCGCGCACCTGCTTTCTTCCACTGTGTACAGACCGTATTTTTC 87007
QY      182 euArgGlyI1eAspValI1eSerProAlaAsnAlaAspThrAspValPheI1eAsnI1eA 202
DB      87008 TGCGGGGATAGACGTTGTTCTCTCTGCCAATGCCGATACAGATGTGTTATTATACATCG 87067
QY      202 spValPheGlyThrI1eArgAsnArgThrGluMetHisLeuTyrAsnAlaGluThrLeuL 222
DB      87068 ACGTAATTCGGAACGATACGAAACGAAACCGAATGCACCTATACAAATGCCAAMACATGA 87127
QY      222 ysaIaGlnThrLysLeuGluTyrPheAlaValAspArgThrAsnLysLysLeuLeuL 242
DB      87128 AAGCCCAAAACAAACTGGAATTTTCGCACTAGACAGAACCAATTAATAATTTGCTCATCA 87187
QY      242 yspProLysThrAsnAlaPheGluAlaI1eTyrLysGluAsnTyrAlaLeuTyrMetGlyP 262
DB      87188 AACCAAAAACCAATGCGTTTGAAGCTGCTATAAAGAAATTAACGATTTGTGAATGGGGC 87247
QY      262 roTyrLysValSerLysGlyI1eLysProThrGluGlyLeuMetValAspPheSerAspI 282
DB      87248 CGTATAAAGTAAGCAAAAGAAATTAAACGACGAAAGATTAATGCTCATTTCTCCGATA 87307
QY      282 leArgProTyrGlyAsnHisThrGlyAsnSerAlaProSerValGluAlaAspAsnSerH 302
DB      87308 TCCGACCATACGGCAATATACGGGTAACTCCGCCCATCCGTAGAGGCTGATTAACAGTTC 87367
QY      302 lsgIuGlyTyrGlyTyrSerAspGlu 310
DB      87368 ATGAGGGGTATGATACAGCATGAA 87393
```

Search completed: August 19, 2005, 03:34:00  
Job time : 605 secs

**This Page Blank (uspto)**